MODERN ALGEBRA

(Abstract Algebra)

[For Degree, Honours and Post-Graduate Students of all Indian Universities and for I. A. S. & P. C. S. Competitions]

By

A.R. VASISHTHA
Department of Mathematics
Meerut College, Meerut.

KRISHNA PRAKASHAN MANDIR
SUBHASH BAZAR, MEERUT-250002, (U.P.) INDIA

Our Most Popular Publications on Mathematics

A. R. Vasishtha
1. Matrices.
2. Modern Algebra.
3. Vector Algebra.

A. R. Vasishtha & D. C. Agarwal
4. Analytical Solid Geometry
5. Dynamics of a Particle.

A. R. Vasishtha & R.K. Gupta
6. Integral Transforms.

D. C. Agarwal
7. Advanced Integral Calculus.
8. Tensor Calculus & Riemannian Geometry.

Dr. Goyal & Gupta
9. Statics.
10. Theory of Relativity.

Dr. Goyal & Sharma
11. Mathematical Statistics.

Dr. Gupta & Malik
12. Calculus of Finite Differences and Numerical Analysis.

Sharma, Gupta & Kumar
13. Spherical Astronomy.

J. N. Sharma
14. Advanced Differential Calculus.
15. Functions of a Complex Variable.
16. Infinite Series & Products.
17. Metric Spaces.
18. Topology.

J. N. Sharma & A. R. Vasishtha
19. Functional Analysis.
20. Linear Algebra.
21. Mathematical Analysis-II.
22. Real Analysis.
23. Vector Calculus.

J. N. Sharma & R. K. Gupta
24. Differential Equations.
25. Mathematical Methods
26. Special Functions.

K. P. Gupta
27. Measure Theory (Analysis-III)
28. Set Theory and Related Topics.

Dr. Mittal & Agarwal
29. Differential Geometry.

Dr. P. P. Gupta.
30. Electro-Magnetic Theory.
31. Rigid Dynamics-I.
32. Rigid Dynamics-II.

R. K. Gupta
33. Linear Programming.
34. Operations Research.

Dr. Shanti Swarup
35. Fluid Dynamics.
36. Hydrodynamics.
37. Linear Integral Equations.

Sharma, Gupta & Agarwal
38. Linear Difference Equations.

MODERN ALGEBRA

First Edition 1960

PREFACE

This book on Modern Abstract Algebra has been written for the use of the students of Honours and Post-graduate classes of Indian Universities.

The subject matter has been discussed in such a simple way that the students will find no difficulty to understand it. The book contains a large number of fully worked out examples. The students should first try to understand the theorems and then they should try to solve the problems independently. Definitions should be read again and again.

Suggestions for the improvement of the book will be gratefully received.

PREFACE TO THE NINETEENTH EDITION

In this edition the book has been thoroughly revised. Suggestions for further improvement of the book will be gratefully received.

—The Author

SYMBOLS

$\exists$	"there exists"
$\forall$	"for every"
$\Rightarrow$	"implies"
$\Leftrightarrow$	"implies and is implied by" or "if and only if" or "iff"
$\wedge$	"and"
$\vee$	"or"
$\in$	"belongs to"
$\notin$	"does not belong to"
$\subseteq$	"is a subset of"
$\supseteq$	"is a superset of"
$\subset$	"is a proper subset of"
$:$ or $\mid$	"such that"
$\cup$	"union"
$\cap$	"intersection"
$\varnothing$	"the null set"
$\mathbf{N}$	"set of natural numbers"
$\mathbf{I}$	"set of integers"
$\mathbf{Q}$	"set of rational numbers"
$\mathbf{R}$	"set of real numbers"
$\mathbf{C}$	"set of complex numbers"
$\mathbf{I}_+$	"set of positive integers"
$\mathbf{R}_+$	"set of positive real numbers"
$\mathbf{Q}_+$	"set of positive rational numbers"
$\mathbf{I}_0$	"set of non-zero integers"
$\mathbf{Q}_0$	"set of non-zero rational numbers"
$\mathbf{R}_0$	"set of non-zero real numbers"
$\mathbf{C}_0$	"set of non-zero complex numbers"

CONTENTS

Chapter *Pages*

1. **Some Basic Set Theoretic Concepts** 1—47
   - Mathematical logic 1
   - Tautologies 2
   - Set 3
   - Subsets of a set 6
   - Union of sets 9
   - Intersection of sets 10
   - Cartesian product of two sets 15
   - Functions or mappings 19
   - Binary operation 33
   - Relations 36
   - Equivalence relations 40
   - Equivalence classes 41
   - Partitions 43
   - Partial order relations 46
2. **Groups** 48—187
   - Binary operation on a set 48
   - Algebraic structure 49
   - Group. Definition 49
   - Abelian group 49
   - Finite and infinite groups 50
   - Order of a finite group 50
   - General properties of groups 55
   - Definition of a group based upon left axioms 58
   - Composition tables for finite sets 69
   - Addition modulo *m* 77
   - Multiplication modulo *p* 77
   - Residue classes of the set of integers 82
   - An alternative set of postulates for a group 88
   - Permutations 93
   - Groups of permutations 96
   - Cyclic permutations 98
   - Even and odd permutations 102
   - Integral powers of an element of a group 111
   - Order of an element of a group 113
   - Isomorphism of groups 126
   - The relation of isomorphism in the set of all groups 133
   - Complexes and subgroups of a group 137
   - Intersection of subgroups 145
   - Cosets 152

Chapter *Pages*
Relation of congruence modulo 158
Lagrange's theorem 159
Euler's theorem 161
Fermat's theorem 162
Order of the product of two subgroups of finite order 162
Cayley's theorem 167
Cyclic groups 170
Subgroup generated by a subset of a group 184
Generating systems of a group 185
3. **Groups (Continued)** **188—253**
Normal subgroups 188
Conjugate elements 197
Normalizer of an element of a group 198
Class equation of a group 200
Centre of a group 200
Conjugate subgroups 203
Invariant subgroups 204
Quotient groups 204
Homomorphism of groups 211
Kernel of a homomorphism 213
Fundamental theorem on homomorphism of groups 216
Automorphisms of a group 221
Inner automorphisms 224
More results on group homomorphism 228
Maximal subgroups 233
Composition series of a group and the Jordan-Holder theorem 234
Solvable groups 236
Commutator subgroup of a group 239
Direct products 244
External direct products 244
Internal direct products 246
Cauchy's theorem on abelian groups 249
Cauchy's theorem 250
Sylow's theorem 251
4. **Rings** **254—353**
Ring. Definition 254
Elementary properties of a ring 255
Rings with or without zero divisors 259
Integral domain 261
Field 262
Division ring or skew field 262
Isomorphism of rings 282
Subrings 286
Subfields 290
Characteristic of a ring 291
Ordered integral domains 294

Chapter *Pages*
Imbedding of a ring into another ring 296
The field of quotients 299
Ideals 306
Principal ideal 319
Principlal ideal ring 320
Divisibility in an integral domain 321
Units 322
Associates 323
Prime elements 325
Greatest common divisor 325
Polynomial rings 325
Polynomials over an integral domain 332
Division algorithm for polynomials over a field 336
Euclidean algorithm for polynomials over a field 339
Unique factorization domain 340
Unique factorization theorem for polynomials over a field 342
Remainder theorem 344
Prime fields 345
Rings of endomorphisms of an abelian group 348
5. Rings (Continued) 354–394
Quotient rings or residue class rings 354
Homomorphism of rings 356
Kernel of a ring homomorphism 357
Maximal ideals 361
Prime ideals 365
Euclidean rings or Euclidean domains 370
Polynomial rings over unique factorization domains 380
6. Vector Spaces. 395–448
Vector space Definition 395
General properties of vector spaces 402
Vector subspaces 403
Linear combination of vectors 410
Linear span 410
Linear sum of two subspaces 412
Linear dependence and linear independence of vectors 413
Basis of a vector space 421
Finite dimensional vector spaces 424
Dimension of a finitely generated vector space 426
Homomorphism of vector spaces or Linear transformation 433
Isomorphism of vector spaces 435
Quotient space 440
Direct sum of spaces 443
Complementary subspaces 446
Co-ordinates 447
7. Vector Spaces (Continued) 449–467
Linear transformations as vectors 449

Chapter *Pages*
Dual space 453
Dual basis 456
Reflexivity 459
Annihilators 463

8. Modules **468—480**
Modules. Definition 468
Submodules 471
Direct sum of submodules 473
Homomorphism of modules or linear transformations 474
Quotient modules 475
Cyclic modules 476
Fundamental theorem on finitely generated modules over Euclidean rings 476

9. Extension Fields and Galois Theory **481—563**
Field extensions 481
Finite field extension 481
Field adjunctions 484
Simple field extension 485
Algebraic field extensions 485
Transcendental element 486
Roots of polynomials 498
Multiple roots 499
Splitting field or decomposition field 503
Uniqueness of the splitting field 512
Derivative of a polynomial 516
Separable extension 523
Perfect field 523
The elements of Galois theory 523
Fixed field 525
Normal extension 529
Galois group 537
Fundamental theorem of Galois theory 539
Construction with ruler and compass 547
Solvability by radicals 552
Finite fields 556
Index *(i-v)*

1

Some Basic Set Theoretic Concepts

§ 1. **Mathematical Logic.** We express our ideas by means of sentences. In mathematics we are concerned only with those sentences which can be judged to be either true or false but not both. Such sentences are called **statements.**

The following are statements :

(*i*) New Delhi is the capital of India (being true).

(*ii*) 8 is greater than 12 (being false).

The following are not statements :

(*i*) How are you ?

(*ii*) The liar says, "I always tell a lie".

We shall use the letters *p, q, r, s* etc. to denote arbitrary statements. We assign to the statement *p* the letter *T* when *p* is true and the letter *F* when it is false. Both *F* and *T* are called the ***truth** values of p.*

Some symbols and notations. The following symbols are very helpful to express our ideas in a compact form :

(*i*) **The symbol $\forall$.** It is used to stand for "*For all*" or "***For** every*". It is called the **universal quantifier.**

For example. $\forall$ real number x, we have $x^2 \geqslant 0$.

(*ii*) **The symbol $\exists$.** It is used to stand for "*There exists.*" It is called the **existential quantifier.**

(*iii*) **The symbol $|$.** It is used to stand for "such that". Sometimes such that is also denoted by : or by s t.

(*iv*) **The symbol $\vee$.** When two or more statements are joined by the word "*or*", the compound statement so formed is called a **disjunction.** The symbol $\vee$ represents the connective "or". Thus the statement $p \vee q$ is read as '*p* or *q*'. The statement $p \vee q$ is true if either *p* or *q* or both *p* and *q* are true. If both *p* and *q* are false, the statement $p \vee q$ is false

(*v*) **The symbol $\wedge$. When two or more statements are joined by the word "*and*", the compound statement so formed is called** a *conjunction.* The symbol used for conjunction is $\wedge$. Thus the statement $p \wedge q$ is read as 'p and q'. The statement $p \wedge q$ is true if and only if both p and q are true. In other words $p \wedge q$ is false if p is false or q is false or both are false.

(*vi*) **The symbol $\Rightarrow$.** If p and q are two statements such that the truth of p implies that of q, we write

$p \Rightarrow q$ (one way implication).

It is read as '*p implies q*'. For example, $x=4 \Rightarrow x^2=16$

We shall assume that the statement $p \Rightarrow q$ is false only when p is true and q is false. In other words a true statement can imply only a true statement while a false statement can imply a true statement or a false statement.

(*vii*) **The symbol $\Leftrightarrow$.** It is used to stand for '*implies and is implied by*' or *if and only if.* Sometimes *iff* is also used to stand for '*if and only if*'. If the truth of the statement p implies that of q and also the truth of q implies that of p we write

$p \Leftrightarrow q$ (both way implication).

For example, $x+2=9 \Leftrightarrow x=7.$

The statement $p \Leftrightarrow q$ is true only when p and q are either both true or both false. It is false when one of the statements is true and the other is false.

(*viii*) **Negation of a Statement.** Associated with each statement is another statement called its **negation**. The negation of a statement p is denoted by '$-p$' or '$\sim p$'. For example, if p is the statement "x is 3", then $\sim p$ is the statement "x is not 3."

The negation of a true statement is false and the negation of a false statement is true.

Tautologies. A statement is called a *tautology* if it is always true. We shall give below some examples of tautologies.

Example 1. *The statement* $(p \wedge q) \Rightarrow p$ is a *tautology.*

We shall prepare the **truth table** for the statement $(p \wedge q) \Rightarrow p$.

p	q	$p \wedge q$	$p \wedge q \Rightarrow p$
T	T	T	T
T	F	F	T
F	T	F	T
F	F	F	T

We observe that the statement $(p \wedge q) \Rightarrow p$ has its truth value T for all its entries in the truth table. Thus it is always true. Hence it is a tautology.

Example 2. *The statement* $\sim(p \wedge q) \Leftrightarrow (\sim p \vee \sim q)$ *is a tautology.*

Truth table for $\sim(p \wedge q) \Leftrightarrow (\sim p \vee \sim q)$

p	q	$p \wedge q$	$\sim(p \wedge q)$	$\sim p$	$\sim q$	$\sim p \vee \sim q$
T	*T*	*T*	*F*	*F*	*F*	*F*
T	*F*	*F*	*T*	*F*	*T*	*T*
F	*T*	*F*	*T*	*T*	*F*	*T*
F	*F*	*F*	*T*	*T*	*T*	*T*

Since the corresponding entries under the columns $\sim(p \wedge q)$ and $\sim p \vee \sim q$ are identical, therefore the statements $\sim(p \wedge q) \Rightarrow (\sim p \vee \sim q)$ and $(\sim p \vee \sim q) \Rightarrow \sim (p \wedge q)$ are both tautologies. Hence the statement $\sim(p \wedge q) \Leftrightarrow (\sim p \vee \sim q)$ is a tautology.

Tautologies are very helpful to us whenever we are to deduce some new statement from some given statements. If the statement $p \Rightarrow q$ is a tautology we can safely conclude the truth of q from the truth of p. We shall now give a list of some important tautologies.

(*i*) $p \vee p \Leftrightarrow p$; $p \wedge p \Leftrightarrow p$. (Idempotent laws)

(*ii*) $p \wedge q \Rightarrow p$; $(p \wedge q) \Rightarrow q$. (Laws of simplification)

(*iii*) $p \Rightarrow (p \vee q)$; $q \Rightarrow (p \vee q)$. (Laws of addition)

(*iv*) $(p \vee q) \Leftrightarrow (q \vee p)$; $(p \wedge q) \Leftrightarrow (q \wedge p)$ (Commutative laws)

(*v*) $\sim(p \wedge q) \Leftrightarrow (\sim p \vee \sim q)$;
$\sim(p \vee q) \Leftrightarrow (\sim p \wedge \sim q)$; (De-Morgan's laws)

(*vi*) $p \wedge (q \wedge r) \Leftrightarrow (p \wedge q) \wedge$
$p \vee (q \vee r) \Leftrightarrow (p \vee q) \vee$ (Associative laws)

(*vii*) $p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$;
$p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$ (Distributive laws)

(*viii*) $p \Leftrightarrow \sim(\sim p)$. (Law of double negation)

(*ix*) $p \vee \sim p$. (Law of the exclusive middle)

(*x*) $(p \Rightarrow q) \Leftrightarrow (\sim q \Rightarrow \sim p)$. (Law of the contrapositive)

(*xi*) $p \wedge (q \wedge r) \Leftrightarrow (p \wedge q) \wedge (p \wedge r)$.

§ 2. A set. Members of a set. (Meerut 1977)

The concept of set is fundamental in all branches of mathematics. *A set is any well-defined class or collection of objects.* By a well defined collection we mean that there exists a rule with the help of which it is possible to tell whether a given object belongs or does not belong to the given collection. The objects in sets may be anything, numbers, people, mountains, rivers etc. The objects constituting the set are called **elements** or **members** of the set.

The following are some of the examples of sets :

(*i*) The numbers 2, 4, 6 and 1.

(*ii*) The countries India, Burma and Afghanistan.

(*iii*) The rivers in India.

(*iv*) The set of all triangles in a plane.

(*v*) The numbers 1, 3, 5, 7,...

We shall denote the sets by capital letters A, B, C, X, Y, Z, etc., and their elements by small letters a, b, c, d, x, y, z etc.

If an object x is a member of a set A, then we write

$$x \in A,$$

which may be read as "*x belongs to A*" or "*x is an element of the set A*". If, on the other hand, an object x is not a member of a set A, then we write

$$x \notin A,$$

which can also be read as "x is not an element of the set A".

It is a common practice in mathematics to put a vertical line "|" or "/" through a symbol to indicate the opposite or negative of that symbol.

A set may be described by actually listing the objects belonging to it For example, let the elements of the set A be a, e, i, o, u; then we write $A=\{a, e, i, o, u\}$. Here the elements are separated by commas and are enclosed in brackets { }. We shall always use these brackets only. This is called the *tabular form* of the set.

A set may also be specified by stating properties which its elements must satisfy. We use some letter, say x, to represent an arbitrary element of the set. The set is then described as follows :

$A=\{x : P(x)\}$ and we say that A is the set consisting of the elements x such that x satisfies the property $P(x)$. The symbol ":" is read "such that". Thus the set A consisting of the elements a, e, i, o, u may be written in this notation as follows :

$A=\{x : x$ is a vowel in the English alphabet$\}$. This way of describing a set is called the **set builder form** of a set.

Finite and Infinite Sets. A set is said to be **finite** if it consists of a specific number of different elements, *i.e.*, if in counting the different members of the set the counting process can come to an end. Otherwise a set is **infinite.**

Example 1. *Let A be the set of the days of the week.* Then A is finite.

Example 2. *Let* $B=\{1, 3, 5, 7, \ldots\}$. Then B is infinite.

§ 3. Some Sets of Numbers. Although in much of our study of sets we shall not be concerned with the type of elements, sets of numbers will naturally appear in most of the examples and problems. We shall now give a list of important sets of numbers along with the symbols we shall often use to denote them.

(*i*) The set $\mathbf{N}=\{1, 2, 3, 4, 5, \ldots\}$ of all natural numbers. The natural numbers are also called counting numbers. The number 0 is not an element of the set of natural numbers.

(*ii*) The set $\mathbf{I}=\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ of all integers. The integers are also called whole numbers.

(*iii*) The set $\mathbf{I}_+=\{1, 2, 3, 4, \ldots\}$ of all positive integers. The number 0 is neither positive nor negative.

(*iv*) The set $\mathbf{Q}=\{x : x=p/q,$ where p and q are integers and $q\neq 0\}$ of all rational numbers.

(*v*) The set $\mathbf{Q}_0$ of all non-zero rational numbers.

(*vi*) The set $\mathbf{R}$ of all real numbers *i.e.*,

$$\mathbf{R}=\{x : x=a\cdot b_1b_2b_3\ldots b_n\ldots\},$$

where $a \in \mathbf{I}$ and b's are any of $0, 1, 2, 3, \ldots, 9$; n can be as large as we please and can go upto infinity. Here every real number has been expressed in the form of decimals with ten as base.

(*vii*) The set $\mathbf{C}$ of all complex numbers, *i.e.*, $\mathbf{C}=\{x : x=a+ib,$ where a and b are real numbers and $i=\sqrt{(-1)}\}$.

§ 4. Equality of Sets. Definition. *Two sets A and B are said to be equal iff every element of A is an element of B and also every element of B is an element of A.* We write "$A=B$" if the sets A and B are equal and "$A\neq B$" if the sets A and B are not equal.

Symbolically, $A=B$ if $x \in A \Leftrightarrow x \in B$.

The statement given in the definition of the equality of two sets is also known as the **axiom of extension.**

Example 1. Let $A=\{1, 2, 3, 4\}$ *and* $B=\{2, 1, 3, 4\}$.

Then by the axiom of extension, $A=B$ since each element of A belongs to B and each element of B belongs to A. Here we see that **a change in the order in which the elements of a set are tabulated is immaterial.**

Example 2. *Let* $A=\{2, 3, 4\}$ *and* $B=\{3, 2, 2, 3, 4\}$.

Then $A=B$ since each element of A is in B and each element of B is in A. Here we see that **a set does not change if one or more of its elements are repeated.** It is for this reason that while describing a set we do not repeat any element. Sometimes we even define a set as "*A well-defined class or collection of distinct objects*".

Example 3. *Let* $A=\{1, 2, 3\}$ *and* $B=\{1, 2, 4, 3, 5\}$.

Then $A \neq B$, since $4 \in B$ but $4 \notin A$.

Example 4 *Let* $A=\{a\}$. Then $a \neq \{a\}$. Here a is an element of the set A. There is a basic difference between an element a and the set $\{a\}$.

§ 5. **Null Set.** *The set which contains no elements at all is called the null set.* This set is sometimes also called the *empty set* or the *void set*. It is denoted by the symbol $\emptyset$.

A set which has at least one element is called a **non-empty set.**

Singleton. A set consisting of a single element is called a singleton. Thus $\{a\}$ is a singleton.

Example 1. *Let* $A=\{x : x^2+1=0$ *and* x *is real*$\}$.

Since there is no real number which satisfies the equation $x^2+1=0$, therefore the set A is empty.

Example 2. *Let* $A=\{x : x$ *is a straight line passing through three distinct points on a circle*$\}$. Then A is the null set.

Example 3. *Let* $A=\{0\}$. Then the set A is not null as it contains one element, *i.e.*, 0. Here A is a singleton. Similarly $\{\emptyset\}$ is not a null set. It is also a singleton.

The null set $\emptyset$ is unique *i.e.*, there exists only one set containing no elements. Therefore the word 'the' is used before an empty set.

§ 6. **Subsets of a set.** **[Meerut 1976]**

If A, B *are sets such that every element of* A *is also an element of* B, *then* A *is said to be a subset of* B. In other words, A is a subset of B *if* $x \in A \Rightarrow x \in B$ For example, if $B=\{x, y, z, t\}$ and $A=\{x, y\}$, then A is a subset of B. Similarly $C=\{x\}$ and $D=\{x, y, t\}$, are also subsets of B. In particular, B itself is also a subset of B.

When A is a subset of B, we denote this relationship by writing

$$A \subseteq B,$$

which is read as '*A is a subset of B*, or '*A is contained in B*'.

If $A \subseteq B$, then B is called a ***super set*** of A and we write

$$B \supseteq A,$$

which is read as '*B is a superset of A*' or '*B contains A*'.

If A is not a subset of B, we write

$$A \nsubseteq B,$$

which is read as '*A is not a subset of B*'.

Similarly $B \nsupseteq A$ is read as 'B is not a super-set of A'.

Proper subsets of a set. From the definition of a subset, it is obvious that *every set is a subset of itself, i.e.,* $A \subseteq A$. If A is any set then A is called an **improper subset** of A. We call B a *proper subset* of A if first, B is a subset of A and, secondly, if B is not equal to A. Briefly, B is a proper subset of A, if $B \subseteq A$ and $B \neq A$.

If B is a proper subset of A, we shall write $B \subset A$.

When we write $B \subseteq A$, then both the possibilities that B is a proper subset of A and $B=A$, are included in it. Some authors, however, denote 'B is a subset of A' by $B \subset A$. Thus in our notation $B \subseteq A$ and in their notation $B \subset A$ will have the same meaning.

If we are to prove that $A \nsubseteq B$, then we should prove that there exists at least one element x such that $x \in A$ but $x \notin B$. Thus the set $A=\{1, 2, 6\}$ is not a subset of $B=\{1, 2, 3, 4, 5\}$, since $6 \in A$ but $6 \notin B$.

Disjoint sets. *If two sets A and B have no elements in common, i.e., if no element of A is in B and no element of B is in A, then we say that A and B are disjoint, or mutually exclusive.* The sets $A=\{1, 3, 7, 8\}$ and $B=\{2, 4, 7, 11\}$ are not disjoint, since the element 7 is common to both A and B. On the other hand, the sets $C=\{a, b, c\}$ and $D=\{x, y, z\}$ are disjoint, since they have no elements in common.

We shall now give some theorems on subsets.

Theorem 1. *If A is any set, then $\varnothing \subseteq A$, i.e., the empty set is a subset of every set.*

Proof. Let us assume that $\varnothing \nsubseteq A$. Then, by our definition of a subset $\exists$ at least one element x such that $x \in \varnothing$ but $x \notin A$. But $\forall\, x$, $x \notin \varnothing$ since $\varnothing$ is the null set. Therefore $\exists$ no element x such that $x \in \varnothing$ and $x \notin A$. Thus our assumption that $\varnothing \nsubseteq A$ is wrong. Hence we must have $\varnothing \subseteq A$.

Theorem 2. *If A and B are sets, then $A=B$ iff $A \subseteq B$ and $B \subseteq A$.*

Proof. If $A=B$, then by the Axiom of Extension every element of A is an element of B and every element of B is an element of A. Hence $A \subseteq B$ and $B \subseteq A$.

Conversely, if $A \subseteq B$ and $B \subseteq A$, then every element of A is an element of B and every element of B is an element of A. Hence, by the axiom of extension, $A=B$.

Theorem 3 *If A is a subset of B and B is a subset of C, then A is a subset of C, that is* $[A \subseteq B, B \subseteq C] \Rightarrow A \subseteq C$.

Proof. Let x be an arbitrary element of the set A. Since $A \subseteq B$, therefore $x \in A \Rightarrow x \in B$. But $B \subseteq C$ (by the hypothesis). Therefore $x \in B \Rightarrow x \in C$. Thus $x \in A \Rightarrow x \in C$. Hence by our definition of a subset $A \subseteq C$.

§ 7. **Class of sets or family of sets.** If the elements of a set are sets themselves, then such a set is said to be a 'class of sets' or a 'family of sets'. In order to avoid confusion we will sometimes use script letters $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ etc. to denote families of sets since capital letters denote their elements.

Example. The set $A=\{\varnothing, \{a\}, \{b\}, \{a, b\}\}$ is a class of sets. Its members are $\varnothing$, $\{a\}$, $\{b\}$, $\{a, b\}$ which are sets themselves.

Important. It should be noted that $\in$ connects an element and a set, while $\subseteq$ connects two sets. Thus if $A=\{r, s, t\}$, then $r \in A$, $\{r\} \subseteq A$ are correct statements, while $r \subseteq A$ and $\{r\} \in A$ are incorrect statements.

§ 8. **Power set.** *If S is any set, then the family of all the subsets of S is called the power set of S.*

The power set of S is denoted by $P(S)$. Symbolically, $P(S)=\{T : T \subseteq S\}$. Obviously $\varnothing$ and S are both elements of $P(S)$.

Example. Let $S=\{a, b, c\}$, then

$P(S)=\{\varnothing, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$.

Theorem. *If a finite set S has n elements, then the power set of S has 2^n elements.*

Proof. Let us form all the subsets of the set S. There are ${}^nC_r$ sub-sets consisting of r elements where $r=0, 1, 2, \ldots, n$.

Since the set S has n elements, therefore no sub-set of S can have more than n elements. Hence the total number of sub-sets of S *i.e.*, the total number of elements in $P(S)$

$$=1+{}^nC_1+{}^nC_2+\ldots+{}^nC_r+\ldots+1=(1+1)^n=2^n.$$

Thus the power set of S has 2^n elements. This is perhaps the motivation for the name power set.

§ 9. **Universal Set.** In any application of set theory, all the sets under consideration will likely be sub-sets of a fixed set. We call this set the *universal set* or *universe of discourse* and denote it by the capital letter U. *Thus a non-empty set of which all the sets under consideration are subsets is called the universal set.*

Example 1. All people in the world constitute the universal set in any study of human population.

Example 2. In plane geometry, the universal set consists of all the points in the plane.

§ 10. **Venn-Euler Diagrams.** We often use pictures in mathematics to help in our thinking. We consider a pictorial representation of sets also. These pictures consist of rectangles and closed curves usually *circles*. These combinations of rectangles and circles are called *Venn-Euler diagrams* or simply, **Venn-diagrams.**

In Venn diagrams a universal set U is represented by a large rectangle and its sub-sets are represented by circular areas drawn within the rectangle. If a set B is a subset of A, the circle representing B is drawn inside the circle representing A. If the sets A and B are equal, then the same circle represents both A and B. If the sets A and B are disjoint, then the circles representing A and B are drawn in such a way that they have no common area. If the sets A and B are not disjoint, the circles representing A and B are drawn in such a way that they have some area common to both.

When we represent a universal set U by a rectangle, we write 'U' in one corner of that rectangle. Similarly we write 'A' within the circle representing the set A.

§ 11. **Union of Sets. Definition.** **(Meerut 1976)**

Let A and B be two sets. The union of A and B is the set of all elements which are in set A or in B. We denote the union of A and B by

$$A \cup B$$

which is usually read as "A union B".

Symbolically, $A \cup B = \{x : x \in A \text{ or } x \in B\}$.

It should be noted here that we take standard mathematical usage of "or". When we say that $x \in A$ or $x \in B$ we do not exclude the possibility that x is a member of both A and B.

Example 1. Let $A=\{a, b, c, d\}$ and $B=\{f, b, d, g\}$.

Then $A \cup B=\{a, b, c, d, f, g\}$.

Example 2. If $A=\{1, 2, 3, 5, 6, 8\}$ and $B=\{0, 2, 4, 7, 9\}$, then $A \cup B=\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

Example 3. If $A=\{a, b, c, d\}$ and $B=\{a, b\}$, *i.e.*, $B \subseteq A$ then $A \cup B=\{a, b, c, d\}=A$.

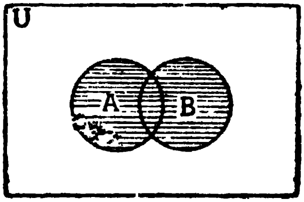
Example 4. If A is the set of all male students in a college and B is the set of all the students in that college, then

$$A \cup B=B.$$

Example 5. If $A=\{1, 3, 5, ...\}$ and $B=\{2, 4, 6, ...\}$, then

$$A \cup B=\{1, 2, 3, 4, 5, 6, ...\}.$$

In the adjoining Venn diagram the union of two sets A and B is shown with shaded area.



§ 12. Intersection of Sets.

Definition. **(Meerut 1976)**

Let A and B be two sets. The intersection of A and B is the set of all elements which are both in A and B. We denote the intersection of A and B by

$$A \cap B$$

which is usually read as "A intersection B". Symbolically,

$$A \cap B=\{x : x \in A \text{ and } x \in B\}$$

or

$$A \cap B=\{x : x \in A, x \in B\}.$$

It should be noted here that the comma has the same meaning as "and".

Example 1. Let $A=\{a, b, c, d\}$ and $B=\{f, b, d, g\}$

Then $A \cap B=\{b, d\}$.

Example 2. Let $A=\{1, 3, 7, 8\}$ and $B=\{2, 4, 9, 11\}$.

Then $A \cap B=\varnothing$, since the sets A and B are disjoint.

Important. *It can be easily seen that two sets A and B are disjoint if and only if* $A \cap B=\varnothing$.

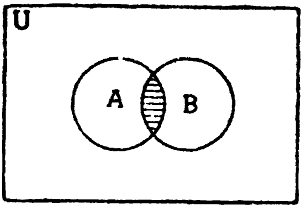
Example 3. If $A=\{1, 2, 4, 6, 8\}$ and $B=\{2, 4, 8\}$ *i.e.*, $B \subseteq A$, then $A \cap B=\{2, 4, 8\}=B$.

Example 4. If A is the set of the students of a college reading chemistry and B is the set of the students of the same college reading mathematics, then $A \cap B$ is the set of the students reading mathematics and chemistry both.

Example 5. If $A=\{x : 0 < x < 3\}$ and $B=\{x : 1 \leqslant x \leqslant 5\}$, then $A \cap B=\{x : 1 \leqslant x < 3\}$.

In the following Venn diagram, the intersection of two sets A and B is shown with shaded area :

$A \cap B$ shaded.



§ 13. **Difference of two Sets. (Meerut 1977).** *The difference of two sets A and B in that order is the set of elements which belong to A but which do not belong to B.*

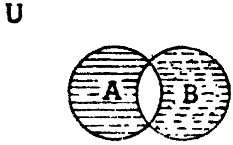
We denote the difference of A and B by

$$A \sim B \text{ or } A - B$$

which is read as "A difference B" or "A minus B". **Symbolically,** $A-B=\{x : x \in A \text{ and } x \notin B\}$. $A-B$ is also called *the complement of B with respect to A.*

Example. Let $A=\{a, b, c, d\}$ and $B=\{f, b, d, g\}$. Then $A-B=\{a, c\}$ and $B-A=\{f, g\}$. Obviously $A-B \neq B-A$.

In the adjoining Venn diagram, area corresponding to $A-B$ has been shaded by horizontal lines and the dotted area represents $B-A$.



Complement of a set relative to the universal set. (Meerut 77)

The complement of a set A with respect to the universal set U is the difference of the universal set U and A. We denote the complement of A with respect to U by A' or by A^c.

Thus $A'=U-A=\{x : x \in U \text{ and } x \notin A\}$ or simply, $A'=\{x : x \notin A\}$ if the universal set is understood.

Obviously $(A')'=A$, $\emptyset'=U$, $U'=\emptyset$.

Example. Let the set of natural numbers $N=\{1, 2, 3, 4, \ldots\}$ be the universal set and let $A=\{1, 3, 5, 7, \ldots\}$. Then

$$A'=\{2, 4, 6, 8, \ldots\}.$$

Symmetric difference of two sets.

Definition. *Suppose A and B are two sets. Then the set $(A-B) \cup (B-A)$ is called the symmetric difference of the sets A and B and is denoted by $A \triangle B$.*

§ **14. Index Set.** Often a non-empty set, say T, has the property, that to each $t \in T$ there corresponds a set C_t. For example, if $T=\{a, b, c, d\}$, then the four corresponding sets are C_a, C_b, C_c, C_d. If $T=\{1, 2, 3, \ldots, n\}$, then the n corresponding sets are C_1, C_2, $\ldots$, C_n. We say that each $t \in T$ *indexes* the sets C_t and when T is used in this sense we call T an *index set.*

Let $\mathbf{A}$ be the collection of sets such that to each member t of T there corresponds a member A_t of $\mathbf{A}$. Then $\mathbf{A}$ is called an indexed family of sets and T is called an index set. Symbolically, we write

$$\mathbf{A}=\{A_t : t \in T\}.$$

§ **15. Arbitrary Union of Sets.** Let T be an index set and suppose that $\forall\ t \in T$, C_t, is a set. The union of the sets C_t, to be denoted by

$$\bigcup_{t \in T} C_t$$

is the set of all elements that are in C_t for some $t \in T$. Symbolically

$$\bigcup_{t \in T} C_t=\{x : x \in C_t \text{ for some } t \in T\}.$$

Example. Let the index set $T=\{a, b, c, d\}$ and let $C_a=\{1, 2\}$, $C_b=\{1, 2, 3, 4\}$, $C_c=\{3, 4, 5\}$, $C_d=\{1, 3, 5, 7\}$.

Then
$$\bigcup_{t \in T} C_t=\{1, 2, 3, 4, 5, 7\}.$$

If $T=\{1, 2, 3, \ldots, n\}$, we may write the union of the sets C_t as follows :

$$\bigcup_{i=1}^{n} C_i=\{x : x \in C_i, \text{ for at least one } i\}.$$

In case the index set T is a singleton, say $T=\{a\}$, then we agree that

$$\bigcup_{t \in T} C_t=C_a.$$

§ **16. Arbitrary intersection of sets.** Let T be an index set and suppose that $\forall\ t \in T$, C_t is a set. The intersection of the sets C_t, to be denoted by

$$\bigcap_{t \in T} C_t$$

is the set of all elements which are in C_t for each $t \in T$.

Symbolically $\bigcap\limits_{t \in T} C_t = \{x : x \in C_t \text{ for each } t \in T\}$.

Example. If $T = \{a, b, c, d\}$ and $C_a = \{1, 2, 3\}$, $C_b = \{2, 3, 4, 5\}$, $C_c = \{2, 3, 5, 7\}$, $C_d = \{1, 2, 3, 5\}$, then $\bigcap\limits_{t \in T} C_t = \{2, 3\}$.

If $T = \{1, 2, 3, \ldots, n\}$, we may write the intersection of the sets C_t as follows :

$$\bigcap_{i=1}^{n} C_i = \{x : x \in C_i \text{ for all } i\text{'s}\}.$$

In case the index set T is a singleton, say $T = \{a\}$, then we agree that $\bigcap\limits_{t \in T} C_t = C_a$.

§ 17. Laws of Algebra of Sets. *If A, B and C are any sets, then*

(*i*) $\left.\begin{array}{l} A \cup B = B \cup A \text{ and} \\ A \cap B = B \cap A \end{array}\right\}$ Commutative laws.

(*ii*) $\left.\begin{array}{l} A \cup (B \cup C) = (A \cup B) \cup C \text{ and} \\ A \cap (B \cap C) = (A \cap B) \cap C. \end{array}\right\}$ Associative laws.

(*iii*) $\left.\begin{array}{l} A \cup A = A \text{ and} \\ A \cap A = A. \end{array}\right\}$ Idempotent laws.

(*iv*) $\left.\begin{array}{l} A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \text{ and} \\ A \cup (B \cap C) = (A \cup B) \cap (A \cup C). \end{array}\right\}$ Distributive laws.

(*v*) $\left.\begin{array}{l} A - (B \cup C) = (A - B) \cap (A - C) \text{ and} \\ A - (B \cap C) = (A - B) \cup (A - C). \end{array}\right\}$ De-Morgan's laws.

(*vi*) If A and B are subsets of a universal set U, then
$\left.\begin{array}{l} (A \cup B)' = A' \cap B' \text{ and} \\ (A \cap B)' = A' \cup B'. \end{array}\right\}$ Another form of De-Morgan's laws.

We shall prove some of these laws. The others may be proved similarly.

1. To prove that $A \cup B = B \cup A$.

Proof. We have $x \in A \cup B \Leftrightarrow x \in A$ or $x \in B$

$\Leftrightarrow x \in B$ or $x \in A$.

[Refer that $p \vee q \Leftrightarrow q \vee p$ is a tautology]

$\Leftrightarrow x \in B \cup A$.

Consequently, $A \cup B \subseteq B \cup A$ and $B \cup A \subseteq A \cup B$.

Hence $A \cup B = B \cup A$.

2. To prove that $A \cap (B \cap C) = (A \cap B) \cap C$.

Proof. We have $x \in A \cap (B \cap C)$

$\Leftrightarrow x \in A$ and $x \in (B \cap C)$

$\Leftrightarrow x \in A$ and ($x \in B$ and $x \in C$)

$\Leftrightarrow$($x \in A$ and $x \in B$) and $x \in C$

[Refer that $p \wedge (q \wedge r) \Leftrightarrow (p \wedge q) \wedge r$ is a tautology]

$\Leftrightarrow x \in A \cap B$ and $x \in C$

$\Leftrightarrow x \in (A \cap B) \cap C.$

Consequently, $A \cap (B \cap C) \subseteq (A \cap B) \cap C$ and

$(A \cap B) \cap C \subseteq A \cap (B \cap C).$

Hence $A \cap (B \cap C) = (A \cap B) \cap C.$

3. To prove that $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ *i.e., the intersection of sets distributes over the union of sets.*

[Dibrugarh 1967; Madras 74]

Proof. We have $x \in A \cap (B \cup C)$

$\Leftrightarrow x \in A$ and $x \in (B \cup C)$

$\Leftrightarrow x \in A$ and ($x \in B$ or $x \in C$)

$\Leftrightarrow$($x \in A$ and $x \in B$) or ($x \in A$ and $x \in C$)

[Refer that $p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$ is a tautology]

$\Leftrightarrow x \in A \cap B$ or $x \in A \cap C$

$\Leftrightarrow x \in (A \cap B) \cup (A \cap C).$

Consequently $A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C)$ and also $(A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C).$

Hence $A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$

4. To prove that $A - (B \cup C) = (A - B) \cap (A - C).$

[Punjab 1976]

Proof. We have $x \in A - (B \cup C)$

$\Leftrightarrow x \in A$ and $x \notin B \cup C$

$\Leftrightarrow x \in A$ and $\sim$($x \in B$ or $x \in C$)

$\Leftrightarrow x \in A$ and ($x \notin B$ and $x \notin C$)

[Refer that $\sim (p \vee q) \Leftrightarrow \sim p \wedge \sim q$ is a tautology]

$\Leftrightarrow$ ($x \in A$ and $x \notin B$) and ($x \in A$ and $x \notin C$)

[Refer that $p \wedge (q \wedge r) \Leftrightarrow (p \wedge q) \Delta (p \wedge r)$ is a tautology]

$\Leftrightarrow x \in (A - B)$ and $x \in (A - C)$

$\Leftrightarrow x \in (A - B) \cap (A - C).$

Hence $A - (B \cup C) = (A - B) \cap (A - C).$

5. To prove that if A **and** B **are subsets of a universal set** U**, then** $(A \cup B)' = A' \cap B'.$ **[Kolhapur 1973]**

Proof. We have

$x \in (A \cup B)' \Leftrightarrow x \notin A \cup B \Leftrightarrow x \notin A$ and $x \notin B$

$\Leftrightarrow x \in A'$ and $x \in B' \Leftrightarrow x \in A' \cap B'.$

Hence $(A \cup B)' = A' \cap B'.$

§ 18. Cartesian Product of two Sets. Let A and B be two given sets. Let $a \in A$ and $b \in B$. Then (a, b) denotes what we may call an **ordered pair.** The object a is called first co-ordinate of the ordered pair (a, b) and the object b is called its second co-ordinate.

If (a, b) and (c, d) are two ordered pairs, then $(a, b)=(c, d)$ iff $a=c$ and $b=d$. Thus the ordered pairs $(2, 5)$ and $(2, 5)$ are equal while the ordered pairs $(2, 5)$ and $(5, 2)$ are different. Here we must note the distinction between the set $\{2, 5\}$ and the ordered pair $(2, 5)$. We have $\{5, 2\}=\{2, 5\}$, but $(2, 5) \neq (5, 2)$.

Definition. *If A and B are sets, the set of all distinct ordered pairs whose first coordinate is an element of A and whose second coordinate is an element of B is called the cartesian product of A and B (in that order) and is denoted by* $A \times B$. *Symbolically,*

$$A \times B=\{(a, b) : a \in A \text{ and } b \in B\}.$$

Example. *Let* $A=\{a, b, c\}$ *and* $B=\{p, q\}$.

Then $A \times B=\{(a, p), (a, q), (b, p), (b, q), (c, p), (c, q)\}$.

Also $B \times A=\{(p, a), (p, b), (p, c), (q, a), (q, b), (q, c)\}$.

We observe that the sets $A \times B$ and $B \times A$ are not equal. Hence in general, $A \times B \neq B \times A$.

If set A has n elements and B has m elements, then the product set $A \times B$ has nm elements.

If either A or B is a null-set, then we agree that $A \times B=\varnothing$. If either A or B is infinite and the other is not empty, then $A \times B$ is infinite.

Product set in General. We may generalise the definition of the product of sets. Let $A_1, A_2, \ldots, A_n$ be n given sets. The set of ordered n-tuples $(a_1, a_2, \ldots, a_n)$ where $a_1 \in A_1, a_2 \in A_2, \ldots\ldots, a_n \in A_n$ is called the cartesian product of $A_1, A_2, \ldots, A_n$ and is denoted by

$$A_1 \times A_2 \times A_3 \times A_4 \times \ldots \times A_n.$$

Solved Examples

Ex. 1. *If A and B are any sets, then prove that*

(*i*) $A \subseteq A \cup B$ *and* $B \subseteq A \cup B$.

(*ii*) $A \subseteq B \Rightarrow A \cup B=B$. (*iii*) $A-B \subseteq A$.

(*iv*) $(A-B) \cup A=A$. (*v*) $(A-B) \cap B=\varnothing$.

Solution. (i) Let x be any arbitrary element of the set A.

Then $x \in A \Rightarrow x \in A$ or $x \in B$

[Refer that $p \Rightarrow p \vee q$ is a tautology]

$\Rightarrow x \in A \cup B$.

Therefore $A \subseteq A \cup B$.

Similarly, we can prove that $B\subseteq A\cup B$.

(ii) Suppose $A\subseteq B$, then to prove that $A\cup B=B$.

We have $x\in A\cup B\Rightarrow x\in A$ or $x\in B$

$\Rightarrow x\in B$ or $x\in B$

[$\because$ $A\subseteq B$, therefore $x\in A\Rightarrow x\in B$]

$\Rightarrow x\in B$. [$\because$ p or $p\Rightarrow p$]

Consequently $A\cup B\subseteq B$.

But $B\subseteq A\cup B$. [see part (i) of this Ex.]

Hence $A\subseteq B\Rightarrow A\cup B=B$.

(iii) Let x be an arbitrary element of the set $A-B$. Then

$x\in A-B\Rightarrow x\in A$ and $x\notin B$

$\Rightarrow x\in A$. [Refer that $p\wedge q\Rightarrow p$ is a tautology]

Consequently $A-B\subseteq A$.

(iv) By part (iii) of this Ex. we have $A-B\subseteq A$.

Therefore by part (ii) of this Ex., we have $(A-B)\cup A=A$.

(v) We have $x\in(A-B)\cap B$

$\Rightarrow x\in A-B$ and $x\in B$

$\Rightarrow (x\in A$ and $x\notin B)$ and $x\in B$

$\Rightarrow x\in A$ and $(x\notin B$ and $x\in B)$

$\Rightarrow x\notin B$ and $x\in B$

[$\because$ $p\wedge q\Rightarrow q$ is a tautology]

But there is no element x which satisfies both $x\notin B$ and $x\in B$. Therefore there is no element in $(A-B)\cap B$,

i.e., $(A-B)\cap B=\varnothing$.

Ex. 2. *If A and B are sets, then prove that $A-B$, $A\cap B$ and $B-A$ are pairwise disjoint.*

[Gorakhpur 1970]

Solution. First we shall prove that $A-B$ and $A\cap B$ are disjoint *i.e.*, $(A-B)\cap(A\cap B)=\varnothing$. We have

$x\in(A-B)\cap(A\cap B)\Rightarrow x\in(A-B)$ and $x\in(A\cap B)$

$\Rightarrow (x\in A$ and $x\notin B)$ and $(x\in A$ and $x\in B)$.

But there is no element x which satisfies both $x\notin B$ and $x\in B$. Therefore there is no element in the set $(A-B)\cap(A\cap B)$ *i.e.*,

$$(A-B)\cap(A\cap B)=\varnothing.$$

Now we shall prove that $B-A$ and $A\cap B$ are disjoint

i.e., $(B-A)\cap(A\cap B)=\varnothing$.

We have $x\in(B-A)\cap(A\cap B)$

$\Rightarrow x\in(B-A)$ and $x\in(A\cap B)$

$\Rightarrow (x\in B$ and $x\notin A)$ and $(x\in A$ and $x\in B)$.

But there is no element x such that $x \notin A$ and $x \in A$. Therefore there is no element in the set

$$(B-A)\cap(A\cap B) \text{ i.e., } (B-A)\cap(A\cap B)=\varnothing.$$

Finally we shall prove that

$$(A-B)\cap(B-A)=\varnothing.$$

We have $x \in (A-B)\cap(B-A)$

$\Rightarrow x \in A-B$ and $x \in B-A$

$\Rightarrow (x \in A$ and $x \notin B)$ and $(x \in B$ and $x \notin A)$.

But there is no element x which satisfies both $x \notin B$ and $x \in B$ or which satisfies both $x \in A$ and $x \notin A$. Therefore $(A-B)\cap(B-A)$ has no element *i.e.*, $(A-B)\cap(B-A)=\varnothing$.

Ex. 3. *Prove that* $(A-B)\cup(B-A)=(A\cup B)-(A\cap B)$.

Solution. We have $x \in (A-B)\cup(B-A)$

$\Leftrightarrow x \in (A-B)$ or $x \in (B-A)$

$\Leftrightarrow (x \in A$ and $x \notin B)$ or $(x \in B$ and $x \notin A)$

$\Leftrightarrow [(x \in A$ and $x \notin B)$ or $x \in B]$ and $[(x \in A$ and $x \notin B)$ or $x \notin A]$

$\Leftrightarrow [x \in B$ or $(x \in A$ and $x \notin B)]$ and $[x \notin A$ or $(x \in A$ and $x \notin B)]$

$\Leftrightarrow [(x \in B$ or $x \in A)$ and $(x \in B$ or $x \notin B)]$ and
$[(x \notin A$ or $x \in A)$ and $(x \notin A$ or $x \notin B)]$

$\Leftrightarrow [(x \in A$ or $x \in B)$ and $(x \in B$ or $x \notin B)]$
and $[(x \notin A$ or $x \in A)$ and $(x \notin A$ or $x \notin B)]$

$\Leftrightarrow x \in A\cup B$ and $x \notin A\cap B \Leftrightarrow x \in (A\cup B)-(A\cap B)$.

Hence $(A-B)\cup(B-A)=(A\cup B)-(A\cap B)$.

Ex. 4. *Prove that* $A\cap(B-C)=(A\cap B)-(A\cap C)$.

Solution. We have $x \in A\cap(B-C)$

$\Rightarrow x \in A$ and $x \in B-C$

$\Rightarrow x \in A$ and $(x \in B$ and $x \notin C)$

$\Rightarrow (x \in A$ and $x \in B)$ and $x \notin C$

$\Rightarrow x \in A\cap B$ and $x \notin C$

$\Rightarrow x \in A\cap B$ and $x \notin A\cap C$

$[\because \; x \notin C \Rightarrow x \notin A\cap C]$

$\Rightarrow x \in (A\cap B)-(A\cap C)$.

$\therefore \; A\cap(B-C)\subseteq(A\cap B)-(A\cap C)$. ...(1)

Again $y \in (A\cap B)-(A\cap C)$

$\Rightarrow y \in A\cap B$ and $y \notin A\cap C$

$\Rightarrow (y \in A$ and $y \in B)$ and $y \notin A\cap C$

$\Rightarrow (y \in A$ and $y \in B)$ and $y \notin C$

$[\because \; y \in A$ and $y \notin A\cap C \Rightarrow y \notin C]$

$\Rightarrow y \in A$ and $(y \in B$ and $y \notin C)$
$\Rightarrow y \in A$ and $y \in (B-C)$
$\Rightarrow y \in A \cap (B-C)$.

$\therefore \quad (A \cap B)-(A \cap C) \subseteq A \cap (B-C)$. ...(2)

From (1) and (2), we get $(A \cap B)-(A \cap C)=A \cap (B-C)$.

Ex. 5. *If* $A \triangle B=(A-B) \cup (B-A)$, *show that*
$A \cap (B \triangle C)=(A \cap B) \triangle (A \cap C)$. **(Banaras 1968)**

Solution. We have $A \cap (B \triangle C)$
$=A \cap [(B-C) \cup (C-B)]$
$=[A \cap (B-C)] \cup [A \cap (C-B)]$ (Distributive Law)
$=[(A \cap B)-(A \cap C)] \cup [(A \cap C)-(A \cap B)]$ (See Ex. 4)
$=(A \cap B) \triangle (A \cap C)$ [by def. of $\triangle$].

Ex. 6. *If A, B and C are sets, prove that*
$A \times (B \cap C)=(A \times B) \cap (A \times C)$.

Solution. Let (x, y) be an arbitrary element of the set $A \times (B \cap C)$.

Then $(x, y) \in A \times (B \cap C)$
$\Leftrightarrow x \in A$ and $y \in B \cap C$
$\Leftrightarrow x \in A$ and $(y \in B$ and $y \in C)$
$\Leftrightarrow (x \in A$ and $y \in B)$ and $(x \in A$ and $y \in C)$
[Refer that $p \wedge (q \wedge r) \Leftrightarrow (p \wedge q) \wedge (p \wedge r)$ is a tautology]
$\Leftrightarrow (x, y) \in (A \times B)$ and $(x, y) \in (A \times C)$
$\Leftrightarrow (x, y) \in (A \times B) \cap (A \times C)$.

Hence $A \times (B \cap C)=(A \times B) \cap (A \times C)$.

Exercises

1. If A and B are any sets, prove that
   (i) $A \cap B \subseteq A$ and $A \cap B \subseteq B$. (ii) $A \subseteq B \Rightarrow A \cap B=A$.
   (iii) $A \cap (A \cup B)=A \cup (A \cap B)=A$.
2. If A and B are subsets of a set X, then prove that
   $A \subseteq B \Leftrightarrow X-B \subseteq X-A$. **(Meerut 1973)**
3. If A, B and C are any sets, prove that
   (i) $A \cup (B \cap C)=(A \cup B) \cap (A \cup C)$ **(Madras 1974)**
   (ii) $A \cap (B-C)=(A \cap B)-C$.
4. Prove that
   (i) $A \triangle B=(A \cup B)-(A \cap B)$, (ii) $A-B=A \triangle (A \cap B)$,
   where $\triangle$ stands for symmetric difference.
5. Prove that $A \subseteq B$ and $C \subseteq D \Rightarrow (A \times C) \subseteq (B \times D)$.
6. Prove that $A \times (B \cup C)=(A \times B) \cup (A \times C)$.
7. If $A=\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$, $B=\{2, 3, 4, 5\}$,
   $C=\{2, 4, 6, 8\}$ and $D=\{4, 5, 6, 7\}$,

find (i) $B \cup C$, (ii) $B \cap D$, (iii) verify that $(B \cup C) \cup (A \cup D) = A$.

(Kolhapur 1973)

8. Which of the following statements are true ? Give reasons to support your answer.

(i) $\{0\} = \varnothing$, where $\varnothing$ denotes the null set.

(ii) $\{s, r, t, r\} = \{r, t, s\}$.

(iii) If $A \subseteq B$, then $B \cup A = A$.

(iv) If $A \subseteq B$, then $A \cap (A - B) = \varnothing$.

(v) $A \cap B = \varnothing$ implies either $A = \varnothing$ or $B = \varnothing$.

Answers

7. (i) $B \cup C = \{2, 3, 4, 5, 6, 8\}$, (ii) $B \cap D = \{4, 5\}$.

8. (i) false, (ii) true, (iii) not true, (iv) true, (v) not true.

§ **19.** **Functions or Mappings.** Let A and B be any two non-empty sets. Let $A = \{a, b, c\}$ and $B = \{x, y, z, t\}$. Suppose by some rule or other, we assign to **each** element of A a **unique** element of B. Let a be associated to x, b be associated to y and c be associated to x. The set $\{(a, x), (b, y), (c, x)\}$ of such assignments is called a **function** from A to B. If we denote this set by f, then we write

$$f : A \to B$$

which is read as "f is a function of A to B or "f is a mapping from A to B".

Function. Definition. **(Kolhapur 1973)**

Let A and B be two given sets. Suppose there exists a correspondence, denoted by f, which associates to each member of A a unique member of B. Then f is called a function or a mapping from A to B. The mapping f of A to B is denoted by

$$f : A \to B \text{ or by } A \xrightarrow{f} B.$$

Range and Domain of a function. Suppose f is a function from A to B. The set A is called the **domain** of the function f, and B is called the **co-domain** of f. The element $y \in B$ which the mapping f associates to an element $x \in A$ is denoted by $f(x)$ and is called the f-image of x or the value of the function f for x. The element x may be referred to as the pre-image of $f(x)$. Each element of A has a unique image and each element of B need not appear as the image of an element in A. There can be more than one elements of A which have the same image in B. We define the **range** of f to consist of those elements in B which appear as the image of at least one element in A. Hence, we often speak of the range of a function as the *image of its domain.* We denote the range of $f: A \to B$ by $f(A)$. Thus $f(A) = \{f(x) : x \in A\}$. Obviously $f(A) \subseteq B$.

Image of a subset. Definition. *Let f be a function of A to B. Let X be any subset of A. Then the set $Y = \{y : y \in B$ and $y = f(x)$ for some $x \in X\}$ is called the image of X under f and we write*

$f(X) = Y$. **(Meerut 1970)**

Functions as sets of ordered pairs. If A and B are any two non-empty sets, then a function f from A to B is a subset f of $A \times B$ satisfying the following conditions :

(*i*) $\forall\, a \in A$, $(a, b) \in f$ for some $b \in B$;

(*ii*) $(a, b) \in f$ and $(a, b') \in f \Rightarrow b = b'$.

The first condition ensures that we have a rule which assigns to each element $a \in A$ some element $b \in B$. Thus each element in A will have image. The second condition guarantees that the image is unique. Accordingly f is a function from A to B.

Example 1. Let $A = \{a, b, c\}$ and $B = \{x, y, z\}$. Then

$$f = \{(a, x), (b, y), (a, z), (c, z)\}$$

is not a function from A to B. The reason is that two elements, x and $z \in B$ are assigned to the same element $a \in A$.

Example 2. Let $A = \{1, 2, 3\}$ and $B = \{1, 2, 3, 4, 5, 6\}$. Let f assign to each member in A its square. Then f is not a function from A to B since no member of B is assigned to the element $3 \in A$.

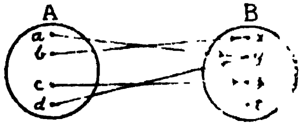
Example 3. Let $A = \{a, b, c, d\}$ and $B = \{a, b, c\}$ Then $f = \{(a, b), (b, c), (c, c), (d, b)\}$ is a function from A to B, since to each element $\in A$ we have assigned a unique element of B. Here $f(a) = b$, $f(b) = c$, $f(c) = c$, and $f(d) = b$. The domain of f is the set A and the range of f *i.e.*, $f(A) = \{b, c\}$.

Transformations or operators. If the domain and co-domain of a function f are both the same set, say $f: A \to A$, then f is often called an **operator** or **Transformation** on A.

Equality of two functions Definition. *Two functions f and g of $A \to B$ are said to be equal iff $f(x) = g(x)$ $\forall\, x \in A$ and we write $f = g$.* For two unequal mappings from A to B, there must exist at least one element $x \in A$ such that $f(x) \neq g(x)$.

Diagrammatic representation of a function. Sometimes a function may be represented by a diagram as will be obvious from the following example.

Let $A=\{a, b, c, d\}$ and $B=\{t, x, y, z\}$. Let $f: A \to B$ be defined by the correspondence $f(a)=y$, $f(b)=x$, $f(c)=z$ and $f(d)=y$.



We represent the two sets A and B by the interiors of two circles. The function $f: A \to B$ is represented by means of a collection of arrows joining the points which represent the elements of A to points representing the corresponding elements of B. By the definition of function, it is obvious that

(*i*) every point in A is joined to some point in B by an arrow;

(*ii*) a point in A cannot be jointed to two or more distinct points in B;

(*iii*) two or more points in A may be joined to the same point in B;

(*iv*) there may be some points in B which are not joined to any point in A.

§ 20. Types of Functions.

'Into' or 'Onto' Mappings. **(Indore 1970; Meerut 77)**

If the mapping $f: A \to B$ is such that there is at least one element in B which is not the f-image of any element in A, then we say that f is a mapping of A 'into' B. In this case the range of f is a proper subset of the co-domain of f i.e., $f(A) \subset B$. Thus in 'into' mappings at least one element of the co-domain B is left uncovered by the f-images of the domain A.

If the mapping $f: A \to B$ is such that each element in B is the f-image of at least one element in A, then we say that f is a mapping of A 'onto' B. In this case the range of f is equal to the co-domain of f i.e., $f(A)=B$. Thus in 'onto' mappings the co-domain B *is completely covered by the f-images of the domain A.* If f is a mapping of A onto B, we write

$$f: A \xrightarrow{\text{onto}} B.$$

Remark. It should be noted that in common use whenever we say that f is a mapping of A into B, we usually include in it the possibility that the mapping f may be onto B also.

'One one' or 'Many-one' mappings. *A mapping $f: A \to B$ is said to be one-one or one-to-one (abbreviated* 1-1) *if different elements in A have different f-images in B i.e., if*

$$f(x)=f(x') \Rightarrow x=x \quad (x \text{ and } x' \in A).$$

In one-one mappings an element in B has only one pre-image in A. If f is a one-one mapping of A to B, we write

$$f: A \xrightarrow[1\text{-}1]{} B.$$

A mapping $f : A \to B$ is said to be many-one if two (or more than two) distinct elements in A have the same f-image in B *i.e.*

$$f(x)=f(x'),\ x \neq x'.$$

In many-one mappings some elements in B have more than one pre-image in A.

One-one onto Mappings. *If $f : A \to B$ is 1-1 and onto B, then f is called a one-to-one correspondence between A and B.*

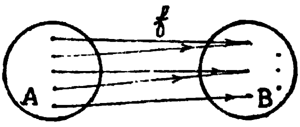
One-one onto mapping is sometimes also known as **bijection.** If $f : A \to B$ is 1-1 and onto B, we write

$$f : A \xrightarrow[1\text{-}1]{\text{onto}} B.$$

Two sets A and B are said to have the same number of elements iff a one-to-one mapping of A onto B exists. Such sets are said to be cardinally equivalent. Cardinally equivalent sets are to have the same cardinal number or the same cardinality.

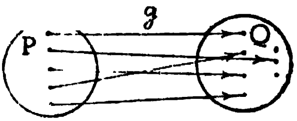
Diagrammatic representation of different types of mappings. **(Meerut 1970)**

1. The mapping $f : A \to B$ in the adjoining diagram is many-one into. Two or more points in A are joined to the same point in B and there are some points in B which are not joined to any point in A.



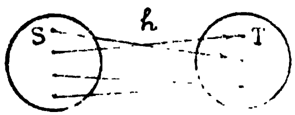
Many-one into

2. The mapping $g : P \to Q$ in the adjoining diagram is one-one into. Different points in P are joined to different points in Q and there are some points in Q which are not joined to any point in P.



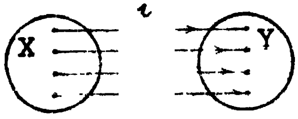
One-one into

3. The mapping $h : S \to T$ in the adjoining diagram is many-one onto. Each point in T is joined to at least one point in S and two or more points in S are joined to the same point in T.



Many-one onto

4. The mapping $i : X \to Y$ in the adjoining diagram is one-one onto. Different points in X are joined to different points in Y and no point in Y is left vacant.



One-one onto

How to prove that a given function f is one-one ?

If we are given a function f and we are to prove that it is one-one, we can do so by showing that if $f(x_1)=f(x_2)$ then $x_1=x_2$, where x_1 and x_2 are arbitrary elements of the domain of f. We can also prove it by showing that if $x_1 \neq x_2$, then $f(x_1) \neq f(x_2)$.

How to prove that a given function $f: A \to B$ is onto ?

To prove that f is onto, we show that if $y \in B$, $\exists\, x \in A$ such that $y=f(x)$. Then $y \in B \Rightarrow y \in f(A)$. Having chosen y arbitrarily, every element of B is an element of $f(A)$ and hence $B \subseteq f(A)$. But $f(A) \subseteq B$. Therefore $B=f(A)$ and the function f is onto B.

Identity Mapping. *Let A be any set. Let the function $f: A \to A$ be defined by the formula $f(x)=x \ \forall\ x \in A$, that is, let each element of A be mapped on itself. Then f is called the identity function, or the identity transformation on A.* We denote this function by I_A. Thus if I_A denotes the identity mapping of a set A, we have $I_A(x)=x \ \forall\ x \in A$.

Example. Let $A=\{1, 2, 3, 4, 5\}$. Then $f=\{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5)\}$ is an identity mapping of A.

Identity mapping is always one-one and onto.

Constant function. *A function $f: A \to B$ is called a constant function if the same element $b \in B$ is assigned to every element in A.* In other words, $f: A \to B$ is a constant function if the range of f consists of only one element.

Example. Let $f: \mathbf{R} \to \mathbf{R}$ be defined by the formula $f(x)=5$. Then f is a constant function since 5 is assigned to every element.

Restriction and extension of a function.

If $f: X \to Y$ and $A \subseteq X$, then the mapping $g: A \to Y$ defined by

$$g(x)=f(x) \ \forall\ x \in A$$

is called the **restriction** *of f to A and is often denoted by $f|A$.*

Also f is called an extension of g.

§ 21. **Inverse image of an element.** Let f be a function of A to B and let $b \in B$. Then the inverse image of the element b under f denoted by $f^{-1}(b)$

consists of those elements in A which have b as their f-image. Symbolically, if $f: A \to B$, then $f^{-1}(b)=\{x : x \in A \text{ and } f(x)=b\}$.

f^{-1} is read as "f inverse". $f^{-1}(b)$ is always a subset of A.

Example 1. Let $A=\{a, b, c, d, e\}$, and $B=\{x, y, z\}$.

Let $f : A \to B$ be defined by
$f=\{(a, x), (b, x), (c, y), (d, x), (e, y)\}$.

Then $f^{-1}(z)$ is the null set $\varnothing$, since no element of A has z as its f-image.

$f^{-1}(y)=\{c, e\}$ since both c and e have y as their f-image.

Also $f^{-1}(x)=\{a, b, d\}$.

Example 2. Let **R** be the set of real numbers and let $f : \mathbf{R} \to \mathbf{R}$ be defined by the formula $f(x)=x^2$. Then $f^{-1}(9)=\{3, -3\}$, since 9 is the f-image of both 3 and -3 and there is no other real number whose square is 9. Also we observe that $f^{-1}(-5)=\varnothing$, since there is no real number whose square is -5.

Inverse image of a subset. Let $f : A \to B$ and let C be a subset of B *i.e.*, $C \subseteq B$. Then the inverse image of C under f, denoted by $f^{-1}(C)$, consists of those elements in A which are mapped into some element in C. Symbolically,

$$f^{-1}(C)=\{x : x \in A \text{ and } f(x) \in C\}.$$

$f^{-1}(C)$ is always a subset of A. In particular $f^{-1}[f(A)]=A$, where $f(A)$ is the range of the function $f : A \to B$. Also $f^{-1}(B)=A$.

§ 22. **Inverse function.** Let $f : A \to B$ be a one-one onto mapping. Let b be any arbitrary element of B. Since the mapping f is onto B, therefore there will be at least one element in A, say a, such that

$$b=f(a),\ b \in B,\ a \in A.$$

Since the mapping f is also one-one, therefore there will be only one element a in A such that $b=f(a)$. Let us denote a by $f^{-1}(b)$. Thus we see that if $f : A \to B$ is one-one onto we can define a new correspondence, denoted by f^{-1}, which associates to each element in B a unique element in A. Accordingly f^{-1} is a function of B to A.

Hence, if $f : A \to B$ is one-one onto, then $f^{-1} : B \to A$. The mapping f^{-1} is called the inverse mapping of the mapping f.

Definition. *Let $f : A \to B$ be a one-one onto mapping. Then the mapping, $f^{-1} : B \to A$, which associates to each element $b \in B$ the element $a \in A$, such that $f(a)=b$, is called the inverse mapping of the mapping $f : A \to B$.*

Only one-one and onto mappings possess inverse mappings. If the mapping $f : A \to B$ is not onto, the correspondence f^{-1} will not be a mapping from B to A. The reason is that in this case there will be some elements in B which will have no f^{-1}-image in A. Again if $f : A \to B$ is not one-one, even then the correspondence f^{-1} will not be a mapping from B to A. The reason is that in this case some elements in B will be associated to more than one element in A. Thus in this case some elements in B will not have a unique f^{-1}-image in A.

Theorem 1. *Let A and B be two sets. If $f: A \to B$ is one-one onto, then $f^{-1}: B \to A$ is also one-one onto.*

Proof. First, let us prove that the mapping f^{-1} is one-one. Let y_1 and y_2 be any two elements of B.

Suppose $f^{-1}(y_1)=x_1$, and $f^{-1}(y_2)=x_2$ where $x_1, x_2 \in A$. Then by the definition of the mapping f^{-1}, $f(x_1)=y_1$, and $f(x_2)=y_2$.

Now $f^{-1}(y_1 = f^{-1}(y_2) \Rightarrow x_1=x_2$

$\Rightarrow f(x_1)=f(x_2)$ [$\because$ f is a mapping from A to B]

$\Rightarrow y_1=y_2$.

$\therefore$ The mapping f^{-1} is one-one.

Now to prove that the mapping f^{-1} is onto A. Let x be any arbitrary element of A. Since f is a mapping from A to B, therefore there exists an element $y \in B$ such that $y=f(x)$ or $x=f^{-1}(y)$. Thus x is the f^{-1}-image of the element $y \in B$. Hence the mapping f^{-1} is also onto.

Theorem 2. *If $f: A \to B$ be one-one and onto, then the inverse mapping of f is unique.*

Proof. Let $f: A \to B$ be one-one and onto. Let $g: B \to A$ and $h: B \to A$ be two inverse mappings of A. To prove that $g=h$.

Let b be any arbitrary element $\in B$. Let $g(b)=a_1$ and $h(b)=a_2$. Since g is an inverse mapping of f, therefore,
$g(b)=a_1 \Rightarrow f(a_1)=b$. Also since h is an inverse mapping of f, therefore $h(b)=a_2 \Rightarrow f(a_2)=b$. But f is a one-one mapping.
$\therefore$ $f(a_1)=b$ and $f(a_2)=b \Rightarrow a_1=a_2 \Rightarrow g(b)=h(b)$. Hence $g=h$.

§ 23. **Intervals defined as Sets of Real Numbers.** Some sets of real numbers occur very frequently. Consider the following sets of real numbers :

$A=\{x : x \in \mathbf{R} \text{ and } 2 \leqslant x \leqslant 7\}$
$B=\{x : x \in \mathbf{R} \text{ and } 2 < x < 7\}$
$C=\{x : x \in \mathbf{R} \text{ and } 2 < x \leqslant 7\}$
$D=\{x : x \in \mathbf{R} \text{ and } 2 \leqslant x < 7\}$.

Such sets are called intervals.

The set A is a closed interval and is denoted by $A=[2, 7]$. The set B is an open interval and is denoted by $B=(2, 7)$. The set C is an open closed interval and is denoted by $C=(2, 7]$. The set D is a closed open interval and is denoted by $D=[2, 7)$.

Solved Examples

Ex. 1. *Decide whether or not the following are functions from A to B where $A=\{1, 2, 3, 4, 5\}$ and $B=\{a, b, c, d, e\}$.*

If they are functions, give the range of each. If they are not, tell, why ?

(*i*) $f=\{(1, a), (2, b), (3, b), (5, e)\}$.

(*ii*) $g=\{(1, e), (5, d), (3, a), (2, b), (1, d), (4, a)\}$.

(*iii*) $h=\{(5, a), (1, e), (4, b), (3, c), (2, d)\}$.

Solution. (i) Since the element $4 \in A$ is not associated to any element $\in B$, therefore f is not a function from A to B.

(ii) The element $1 \in A$ is associated to two different elements e and $d \in B$. Therefore g is not a function from A to B.

(iii) Each element of A is associated to a unique element of B. Therefore h is a function from $A \rightarrow B$. The range of h is the set of the h-images of all elements of A. So range of $h=h(A)$ $=\{a, e, b, c, d\}=B$.

Ex. 2. *Let $A=\{-2, -1, 0, 1, 2\}$. Let the function $f: A \rightarrow \mathbf{R}$ be defined by the formula $f(x)=x^2+1$. Find the range of f.*

Solution. The range of f consists of those elements of $\mathbf{R}$ which appear as f-images of different elements of A. So we calculate the f-image of each element of A.

$f(-2)=(-2)^2+1=5$, $f(-1)=(-1)^2+1=2$,

$f(0)=(0)^2+1=1$, $f(1)=(1)^2+1=1+1=2$,

$f(2)=(2)^2+1=4+1=5$.

Thus the range of f is the set $\{5, 2, 1, 2, 5\}$, *i.e.*, the set $\{5, 2, 1\}$.

Ex. 3. *Each of the following formulas defines a function from* $\mathbf{R}$ *to* $\mathbf{R}$. *Find the range of each function.*

(*i*) $f(x)=x^3$, (*ii*) $g(x)=\sin x$, (*iii*) $h(x)=x^2+1$.

Solution. (i) We know that every real number a has a real cube root $\sqrt[3]{(a)}$. Therefore if a be any arbitrary element $\in \mathbf{R}$, then $f(\sqrt[3]{a})=(\sqrt[3]{a})^3=a$ *i.e.*, $\sqrt[3]{a} \in \mathbf{R}$ is the pre-image of $a \in \mathbf{R}$. Since a is any arbitrary element $\in \mathbf{R}$, therefore the range of f is $\mathbf{R}$. This function will be an onto function.

(ii) The sine of any real number lies in the closed interval $[-1, 1]$. Also, all the numbers in this interval will be the sine of some real numbers. Hence the range of g is the closed interval $[-1, 1]$. This function will be an into function.

(iii) If we add 1 to the square of each real number, we get the set of numbers which are greater than or equal to 1. Let y be any real number greater than or equal to 1 and let $y=h(x)=x^2+1$. Then $x=\pm\sqrt{(y-1)}$ are also real numbers.

Thus every real number which is greater than or equal to 1 is the h-image of some or other real number. Hence the range of h is the infinite interval $[1, \infty)$.

Ex. 4. *Let* Q *be the set of rational numbers. Let* $f: Q \to Q$ *be defined by* $f(x)=2x+3$, $(x \in Q)$. *Show that* f *is one-one onto. Also find a formula that defines the inverse function* f^{-1}.

Solution. Let m and n be any two different elements in Q. Then $m \neq n \Rightarrow 2m \neq 2n \Rightarrow 2m+3 \neq 2n+3 \Rightarrow f(m) \neq f(n)$.

Thus different elements in Q have different f-images in Q. Hence f is one-one.

Let y be any arbitrary element in Q. If $y=f(x)=2x+3$, we have $x=(y-3)/2$ which is also a rational number.

Thus $f\left(\frac{y-3}{2}\right)=y$ *i.e.*, any arbitrary element y in Q is the f-image of the element $(y-3)/2 \in Q$. Hence f is onto.

Since $f: Q \to Q$ is one-one onto, therefore f has an inverse function $f^{-1}: Q \to Q$.

Let y be the image of x under the function f. Then $y=f(x)=2x+3$. Consequently, x will be the image of y under the inverse function f^{-1} *i.e.*, $x=f^{-1}(y)$. Solving for x in terms of y in the equation $y=2x+3$, we get $x=(y-3)/2$.

Thus $f^{-1}(y)=\frac{y-3}{2}$, $(y \in Q)$ is the formula defining the inverse function $f^{-1}: Q \to Q$.

Note. In order to prove that the mapping f is one-one, we can also argue like this : Let m and n be any two elements in Q. Then $f(m)=f(n) \Rightarrow 2m+3=2n+3 \Rightarrow 2m=2n \Rightarrow m=n$. Therefore f is one-one.

Ex. 5. *Let* $X=\{x : x \in R$ *and* $-\pi/2 \leqslant x \leqslant \pi/2\}$ *i.e*, *let*
$X=[-\pi/2, \pi/2]$ *and*
$Y=\{y : y \in R$ *and* $-1 \leqslant y \leqslant 1\}$ *i.e., let* $Y=[-1, 1]$.
Show that the function $f: X \to Y$ *defined by*
$f(x)=\sin x$, $(x \in X)$, *is one-one onto.*
Also give the inverse map $f^{-1}: Y \to X$.

Solution Let m and n be any two different real numbers lying in the closed interval $[-\pi/2, \pi/2]$. We know that any two different real numbers lying in the closed interval $[-\pi/2, \pi/2]$ have not the same sine.

$\therefore$ $m \neq n \Rightarrow \sin m \neq \sin n \Rightarrow f(m) \neq f(n)$. Hence f is one-one.

Again if y is any arbitrary real number lying in the closed interval $[-1, 1]$, $\exists$ a real number x lying in the closed interval $[-\pi/2, \pi/2]$ such that $\sin x = y$.

Thus every element y in Y is the f-image of some element x in X. Hence f is onto.

Thus $f : X \to Y$ is one-one onto, therefore f has an inverse function $f^{-1} : Y \to X$.

Let y be the image of x under the function f. Then $y = f(x) = \sin x$. Consequently, x will be the image of y under the inverse function f^{-1} *i.e.*, $x = f^{-1}(y)$. Solving for x in terms of y, in the equation $y = \sin x$, we get $x = \sin^{-1} y$. Thus $f^{-1}(y) = \sin^{-1} y$, $(y \in Y)$ is the formula defining the inverse function $f^{-1} : Y \to X$.

Ex. 6. *Let* **C** *be the set of complex numbers. Prove that the map* $f : \mathbf{C} \to \mathbf{R}$ *given by* $f(z) = |z|$, $z \in \mathbf{C}$ *is neither one-one nor onto.*

Solution. Let $z = x + iy$ be any complex number, where x, $y \in \mathbf{R}$ and $i = \sqrt{(-1)}$. Then $|z| = \sqrt{(x^2 + y^2)}$. Also $|z|$ is always a non-negative real number *i. e.*, $|z| \geqslant 0$.

If m be any negative real number $\in \mathbf{R}$, then there exists no complex number $z \in \mathbf{C}$ such that $|z| = m$. Thus m is not the f-image of any complex number $\in \mathbf{C}$. Hence f is not onto but is into.

Also $z_1 = 2 + i3 \in \mathbf{C}$ and $z_2 = 2 - i3 \in \mathbf{C}$. Then $|2 + i3| = \sqrt{13}$ and $|2 - i3| = \sqrt{13}$. Thus $z_1 \neq z_2$, although $|z_1| = |z_2|$ *i.e.*, $f(z_1) = f(z_2)$. Thus we see that two different complex numbers z_1 and $z_2 \in \mathbf{C}$ have the same f-image in **R**. Hence f is not one-one.

Ex. 7. *If S and T are non-empty sets, prove that there exists a one-to-one correspondence between $S \times T$ and $T \times S$.*

Solution. If $a \in S$ and $b \in T$, then the ordered pair $(a, b) \in S \times T$ and the ordered pair $(b, a) \in T \times S$. Let f be a mapping from $S \times T$ to $T \times S$ defined by the formula

$$f(a, b) = (b, a) \ \forall \ (a, b) \in S \times T.$$

Here $f(a, b)$ denotes the image under the mapping f of the element $(a, b) \in S \times T$.

Obviously the mapping f is well-defined.

f is one-one. Let $(a, b), (c, d) \in S \times T$. Then $f(a, b) = f(c, d)$
$\Rightarrow (b, a) = (d, c)$ [by def. of f]
$\Rightarrow b = d, a = c \Rightarrow (a, b) = (c, d) \Rightarrow f$ is one-one.

f is onto. Let (y, z) be any element of $T \times S$. Then $y \in T$, $z \in S$. Therefore $(z, y) \in S \times T$. We have $f(z, y) = (y, z)$.

Thus $(y, z) \in T \times S \Rightarrow \exists\ (z, y) \in S \times T$ such that $f(z, y) = (y, z)$. Therefore f is onto.

Thus f is a one-one function from $S \times T$ onto $T \times S$. Therefore f gives a one-to-one correspondence between $S \times T$ and $T \times S$.

Ex. 8. *Let* **R** *be the set of all real numbers. Using the fact that every cubic equation with real coefficients has a real root, show that* $x \to x^3 - x$ *defines a mapping of* **R** *onto* **R**. *Is this a one-one mapping ?*

Solution. If $x \in$ **R**, then $x^3 - x \in$ **R** and is unique. Therefore $x \to x^3 - x$ defines a mapping of **R** to **R**.

Let y be any arbitrary element $\in$ **R** *i.e.*, let y be any real number. Then $x^3 - x = y$ is a cubic equation with real coefficients. It will have at least one real root. Thus for any $y \in$ **R** there exists $x \in$ **R** such that $x^3 - x = y$. Therefore $x \to x^3 - x$ defines a mapping of **R** onto **R**.

Again $1 \to 1^3 - 1$ *i.e.*, $1 \to 0$ and $-1 \to (-1)^3 - (-1)$ *i.e.*, $-1 \to 0$. Thus the two elements 1 and $-1 \in$ **R** map onto the same element $0 \in$ **R**. Hence the mapping is not one-one.

§ 24. Product or Composite of mappings.

Let f be a function of X to Y and let g be a function of Y to Z. Here the domain of the function g is the co-domain of the function f. Let x be any arbitrary element in X *i.e.*, let $x \in X$. Then the image of x under f *i.e.*, $f(x)$ is in Y. Since $g : Y \to Z$ and $f(x) \in Y$, therefore we can find the image of $f(x)$ under g *i.e.*, we can find $g\ [f(x)]$ which will be in Z. Also $f(x)$ is unique and consequently $g\ [f(x)]$ is also unique. Thus we have a rule which assigns to each element $x \in X$ a unique element $g\ [f(x)] \in Z$. In this way we have a function of X to Z. This new function is called the *product function* or *composite function* of f and g and it is denoted by $(g \circ f)$ or (gf).

Definition. *Let* $f : X \to Y$ *and* $g : Y \to Z$.

Then the composite of the functions f *and* g *denoted by* $(g \circ f)$, *is a mapping of* $X \to Z$ *given by*

$$(g \circ f) : X \to Z \text{ such that}$$

$(g \circ f)(x) = g\ [f(x)],\ \forall\ x \in X$. **[Jabalpur 1970]**

Example 1. Let $X = \{1, 2, 3, 4\}$, $Y = \{a, b, c, d\}$ and $Z = \{l, m, n\}$. Let $f : X \to Y$ and $g : Y \to Z$ be defined by $f = \{(1, a), (2, c), (3, b), (4, a)\}$ and $g = \{(a, l), (b, l), (c, m), (d, m)\}$. Then the composite mapping $(g \circ f) : X \to Z$ may be computed in the following manner :

$(g \circ f)(1) = g\ [f(1)] = g\ (a) = l$, $(g \circ f)(2) = g\ [f(2)] = g\ (c) = m$,

$(g \circ f)(3)=g[f(3)]=g(b)=l$, $(g \circ f)(4)=g[f(4)]=g(a)=l$.

Thus $(g \circ f): X \to Z$ is given by

$$g \circ f=\{(1, l), (2, m), (3, l), (4, l)\}.$$

Example 2. Let $f: \mathbf{R} \to \mathbf{R}$ be given by the formula

$$f(x)=\sin x, (x \in \mathbf{R})$$

and the map $g: \mathbf{R} \to \mathbf{R}$ be given by the formula

$$g(x)=x^2, (x \in \mathbf{R})$$

The composite map $(g \circ f): \mathbf{R} \to \mathbf{R}$ is given by the formula $(g \circ f)(x)=g[f(x)]=g(\sin x)=(\sin x)^2=\sin^2 x, (x \in \mathbf{R})$.

Note. If $f: A \to A$ and $g: A \to A$, then we can find both the composite mappings $g \circ f$ and $f \circ g$. But in general $(g \circ f) \neq (f \circ g)$. As in example 2, the composite map $(f \circ g): \mathbf{R} \to \mathbf{R}$ is given by the formula $(f \circ g)(x)=f[g(x)]=f(x^2)=\sin x^2, (x \in \mathbf{R})$.

We see that $(f \circ g): \mathbf{R} \to \mathbf{R}$ and $(g \circ f): \mathbf{R} \to \mathbf{R}$ are defined by different formulas.

§ 25. Some Properties of Composites of Mappings.

Theorem 1. *If $f: X \to Y$ be a one-one and onto mapping, then*

$$f \circ f^{-1}=I_Y \text{ and } f^{-1} \circ f=I_X,$$

where I_X and I_Y are the identity mappings of the sets X and Y respectively. **(Allahabad 1980)**

Proof. We shall first prove that $f^{-1} \circ f=I_X$.

Since $f: X \to Y$ and $f^{-1}: Y \to X$, therefore $(f^{-1} \circ f): X \to X$.

Let x be any arbitrary element $\in X$ and let $f(x)=y$, where $y \in Y$. Then by the definition of the mapping f^{-1}, $f^{-1}(y)=x$.

We have $(f^{-1} \circ f)(x)=f^{-1}[f(x)]=f^{-1}(y)=x$. Thus the function $f^{-1} \circ f$ maps every element $x \in X$ onto itself. Therefore $f^{-1} \circ f=I_X$.

Now we shall prove that $f \circ f^{-1}=I_Y$. Since $f^{-1}: Y \to X$ and $f: X \to Y$, therefore $(f \circ f^{-1}): Y \to Y$.

Let y be any arbitrary element $\in Y$ and let $f^{-1}(y)=x$, where $x \in X$. Then $f(x)=y$

We have $(f \circ f^{-1})(y)=f[f^{-1}(y)]=f(x)=y$. Thus the function $f \circ f^{-1}$ maps every element $y \in Y$ onto itself. Therefore $f \circ f^{-1}=I_Y$.

Theorem 2. *If $f: X \to Y$, then $I_Y \circ f=f$ and $f \circ I_X=f$ i.e., the product of any function with the identity function is the function itself.*

Proof. Let x be any arbitrary element $\in X$ and let

$$f(x)=y, (x \in X, y \in Y).$$

Since $f: X \to Y$ and $I_Y: Y \to Y$, therefore $(I_Y \circ f): X \to Y$.

We have $(I_Y \circ f)(x)=I_Y[f(x)]=I_Y(y)=y$.

[$\because$ I_Y is the identity mapping of Y]

Also $f(x)=y$.

Thus $\forall\ x \in X$, $(I_Y \circ f)(x)=f(x)$. Therefore $I_Y \circ f=f$.

Again since $I_X : X \to X$ and $f : X \to Y$, therefore $f \circ I_X : X \to Y$.

We have $(f \circ I_X)(x)=f[I_X(x)]=f(x)$

[$\because$ I_X is the identity mapping of X]

Thus $\forall\ x \in X$, $(f \circ I_X)(x)=f(x)$. Therefore $f \circ I_X=f$.

Theorem 3. *If $f : X \to Y$ and $g : Y \to Z$ be two one-one onto maps, then $g \circ f : X \to Z$ is also one-one onto and*

$(g \circ f)^{-1} : Z \to X=(f^{-1} \circ g^{-1}) : Z \to X$. **(Punjab 1970)**

Proof. We shall prove that

(i) $g \circ f$ is one-one, (ii) $g \circ f$ is onto, (iii) $(g \circ f)^{-1}=f^{-1} \circ g^{-1}$.

(i) Let x_1 and x_2 be any two elements in X. The mapping $g \circ f$ will be one-one if we show that the images of x_1 and x_2 under this mapping are equal only if $x_1=x_2$. We have

$(g \circ f)(x_1)=(g \circ f)(x_2) \Rightarrow g[f(x_1)]=g[f(x_2)]$

[by the definition of the composite mapping $g \circ f$]

$\Rightarrow f(x_1)=f(x_2)$ [$\because$ $f(x_1), f(x_2) \in Y$ and the mapping $g : Y \to Z$ is one-one]

$\Rightarrow x_1=x_2$ [$\because$ f is one-one]

$\therefore$ the mapping $g \circ f$ is one-one.

(ii) Let z be any arbitrary element of Z. Since g is an onto mapping of Y to Z, therefore $\exists\ y \in Y$, such that $g(y)=z$. Again, since f is an onto mapping of X to Y, therefore $\exists\ x \in X$, such that $f(x)=y$. Thus for every $z \in Z$, $\exists\ x \in X$, such that $z=g(y)$ $=g[f(x)]=(g \circ f)(x)$. Therefore $g \circ f$ is a mapping of X onto Z.

Thus $g \circ f$ is one-one onto and hence is inversible.

(iii) Let z be any arbitrary element $\in Z$, such that $z=g(y)$, $(y \in Y, z \in Z)$. Then $y=g^{-1}(z)$.

Also let $y=f(x)$, $(y \in Y, x \in X)$. Then $x=f^{-1}(y)$.

We have $(g \circ f)(x)=g[f(x)]=g(y)=z$.

Since $g \circ f$ is one-one onto,

$\therefore$ $(g \circ f)(x)=z \Rightarrow (g \circ f)^{-1}(z)=x$.

Also $(f^{-1} \circ g^{-1})(z)=f^{-1}[g^{-1}(z)]=f^{-1}(y)=x$.

Thus $\forall\ z \in Z$, $(g \circ f)^{-1}(z)=(f^{-1} \circ g^{-1})(z)$.

Hence the mappings $(g \circ f)^{-1}$ and $f^{-1} \circ g^{-1}$ are equal.

Theorem 4. Associativity of Composites of Functions.

Let $f : A \to B$, $g : B \to C$ and $h : C \to D$, then

$(h \circ g) \circ f=h \circ (g \circ f)$. **(Jaipur 1970)**

Proof. It can be easily seen that both $(h \circ g) \circ f$ and $h \circ (g \circ f)$ are mappings of A to C. These two mappings will be equal if they assign the same image to each element x in the domain A, *i.e.*, if

$$[(h \circ g) \circ f]\,(x)=[h \circ (g \circ f)]\,(x).$$

We have, by definition of composites of mappings,

$$[(h \circ g) \circ f]\,(x)=(h \circ g)\,[f(x)]=h\,[g\,\{f(x)\}]=h\,[(g \circ f)\,(x)]$$
$$=[h \circ (g \circ f)]\,(x).$$

Hence $(h \circ g) \circ f=h \circ (g \circ f)$.

Theorem 5. *Let $f : X \to Y$ and $g : Y \to X$. Then g is the inverse function of f i.e., $g=f^{-1}$, if the product function $(g \circ f) : X \to X$ is identity function on X and $(f \circ g) : Y \to Y$ is the identity function on Y.* **(Punjab 1970)**

Proof. In order to show that f is invertible, we are to show that f is one-one and onto.

(i) f **is one-one.** Let $x_1, x_2 \in X$. Then

$$\begin{aligned} f(x_1)=f(x_2) &\Rightarrow g\,[f(x_1)]=g\,[f(x_2)] \\ &\Rightarrow (g \circ f)\,(x_1)=(g \circ f)\,(x_2) \\ &\Rightarrow I\,(x_1)=I\,(x_2) \qquad [\because\ g \circ f=I] \\ &\Rightarrow x_1=x_2. \end{aligned}$$

$\therefore$ f is one-one.

(ii) f **is onto.** Let y be any element of Y. Since g is a function from Y to X, therefore $g\,(y) \in X$.

Let $g\,(y)=x$. Then

$$\begin{aligned} g\,(y)=x &\Rightarrow f\,[g\,(y)]=f\,(x) \\ &\Rightarrow (f \circ g)\,(y)=f\,(x) \\ &\Rightarrow I\,(y)=f\,(x) \qquad [\because\ f \circ g=I] \\ &\Rightarrow y=f\,(x). \end{aligned}$$

Thus $y \in Y \Rightarrow \exists\, x \in X$ such that $f\,(x)=y$. Therefore f is onto. Since f is 1-1 onto, therefore f is invertible *i.e.*, f^{-1} exists.

(iii) Now we shall prove that $f^{-1}=g$. We have

$$f \circ g=I \Rightarrow f^{-1} \circ (f \circ g)=f^{-1} \circ I$$
$$\Rightarrow (f^{-1} \circ f) \circ g=f^{-1} \Rightarrow I \circ g=f^{-1} \Rightarrow g=f^{-1}.$$

Note. In a similar manner we can show that g is also invertible and $g^{-1}=f$.

Exercises

1. Define a 'function' giving examples, and distinguish between 'onto' and 'into' mappings. Give an example of a 'one-to-one correspondence' after defining it.

Let $\mathbf{R}_0$ denote the set of all non-zero real numbers. Prove that the map $f : \mathbf{R}_0 \to \mathbf{R}_0$ given by $f(x)=1/x$, $x \in \mathbf{R}_0$ is both one-one and onto. (Meerut 1977)

2. Explain what do you understand by the domain and the range of a mapping. Obtain the domain and the range of the mapping

$$f : \mathbf{R} \to \mathbf{R} : f(x) = \sin x$$

where **R** is the set of real numbers. **(Meerut 1976)**

3. If **R** is the set of all real numbers, show that $\phi : x \to x^3 + x^2$ defines a mapping of **R** onto **R**. Is ϕ one-one ? Justify your answer.

4. Show that the mapping $f : \mathbf{R} \to \mathbf{R}$ defined by

$$f(x) = \cos x, \ \forall \ x \in \mathbf{R}$$

is neither one-one nor onto. Modify the domain and co-domain of this mapping so that it may be both one-one and onto.

5. Prove the following two results for a finite set S :

(i) If f maps S onto S then f is one-to-one.

(ii) If f is a one-one mapping of S into itself, then f is onto.

Prove, by example, that both (i) and (ii) are false if S is a set not having a finite number of elements.

6. Prove that the mapping $f : A \to B$ is one-one onto if and only if there exists a mapping $g : B \to A$ such that $g \circ f$ and $f \circ g$ are the identity mappings on A and B respectively.

7. If the mapping $f : \mathbf{R} \to \mathbf{R}$ be given by $f(x) = 4x - 1$, and the mapping $g : \mathbf{R} \to \mathbf{R}$ be given by $g(x) = x^3 + 2$, then find formulae defining $g \circ f$ and $f \circ g$ **(Kolhapur 1973)**

8. If f and g are two mappings from **R** to **R** given by

$$f(x) = x^2 + 3x + 1, \ g(x) = 2x - 3,$$

then obtain formulae defining $f \circ g$ and $g \circ g$. **(Meerut 1973)**

9. Give an example to show that composition of two functions is not in general commutative. **(Meerut 1974)**

Answers

2. Domain $f = \mathbf{R}$, range $f =$ the closed interval $[-1, 1]$.

3. ϕ is not one-one. We have $\phi(0) = 0$, $\phi(-1) = 0$.

4. Take $X = \{x : x \in \mathbf{R} \text{ and } -\pi/2 \leqslant x \leqslant \pi/2\}$ as the domain of f and $Y = \{y : y \in \mathbf{R} \text{ and } -1 \leqslant y \leqslant 1\}$ as the co-domain of f.

7. $(g \circ f)(x) = 64x^3 - 48x^2 + 12x + 1$, $(f \circ g)\, x = 4x^3 + 7$.

8. $(f \circ g)\, x = 4x^2 - 6x + 1$, $(g \circ g)\, x = 4x - 9$.

§ 26 **Binary Operation.** *A binary operation* "o" *on a non-empty set A is a mapping which associates with each ordered pair (a, b) of elements of A, a uniquely defined element $c \in A$. Thus* "o" is a mapping of the product set $A \times A$ to A. Symbolically, a map $\circ : A \times A \to A$, is called a binary operation on the set A.

The o-image of the element $(a, b) \in A \times A$ is denoted by $a \circ b$.

If a set A is closed with respect to the composition o, then we say that o is a binary operation on the set A.

Example 1. The composition of addition is a binary operation on the set of even natural numbers. The sum of two even natural numbers is also an even natural number. Thus the set of even natural numbers is closed under the composition of addition. On the other hand, addition is not a binary operation on the set of odd natural numbers. The sum of two odd natural numbers is an even natural number. Thus the sum of two odd natural numbers does not belong to the set of odd natural numbers. In this way the set of odd natural numbers is not closed under addition.

Example 2. The set of natural numbers N is not closed under subtraction. For example, $4 \in \mathbf{N}$ and $3 \in \mathbf{N}$, but $3-4 \notin \mathbf{N}$. Thus subtraction is not a binary operation on N. On the other hand, the students may verify that subtraction is a binary operation on I.

§ 27. Types of Binary Operations.

Commutative Operations. *A binary operation* o *on a set* A *is called commutative if* $a \circ b = b \circ a$, *for every* $a, b \in A$.

Example. Addition and multiplication are commutative binary operations on the set of real numbers since

$$a+b=b+a \text{ and } ab=ba, \text{ (for all } a, b \in \mathbf{R}).$$

Associative Operations. *A binary operation* o *on a set* A *is called associative if* $(a \circ b) \circ c = a \circ (b \circ c)$, *(for every* $a, b, c \in A$*).*

Example 1. Addition and multiplication are associative binary operations on N, since

$$a+(b+c)=(a+b)+c \text{ and } a(bc)=(ab)c, \text{ (for every } a, b, c \in \mathbf{N}).$$

Example 2. Let o be a binary operation on **R** defined by

$$a \circ b = a+2b \text{ for all } a, b \in \mathbf{R}.$$

Then $(a \circ b) \circ c = (a+2b) \circ c = a+2b+2c$,

and $a \circ (b \circ c) = a \circ (b+2c) = a+2(b+2c) = a+2b+4c$.

Thus $(a \circ b) \circ c \neq a \circ (b \circ c)$, for every $a, b, c \in \mathbf{R}$.

Hence o is not an associative binary operation on **R**.

Distributive Operations. *Let* A *be a set on which two binary operations denoted by* o *and* $*$ *are defined. The operation* $*$ *is said to be left distributive with respect to* o *if*

$$a * (b \circ c) = (a * b) \circ (a * c), \text{ for all } a, b, c \in A \quad \text{...(1)}$$

and is said to be right distributive with respect to o *if*

$$(b \circ c) * a = (b * a) \circ (c * a), \text{ for all } a, b, c \in A. \quad \text{...(2)}$$

When both (1) and (2) hold, we say simply that $*$ is distributive with respect to o.

If $\circ$ is commutative, then (2) will hold if (1) holds and vice-versa

Example. For the set **R** of all real numbers, multiplication $(* = \cdot)$ is distributive with respect to addition $(\circ = +)$, since

$a \cdot (b+c) = a \cdot b + a \cdot c$ and $(b+c) \cdot a = b \cdot a + c \cdot a \ \forall\ a, b, c \in \mathbf{R}$.

But addition is not distributive with respect to multiplication, since $a + (b \cdot c) \neq (a+b) \cdot (a+c)$, for every $a, b, c \in \mathbf{R}$.

Identity Element. *Let* $\circ : A \times A \to A$ *be a binary operation on* A. *An element* $e \in A$ *is called an identity element for the operation* $\circ$ *if*

$$e \circ a = a \circ e = a, \text{ for every } a \in A.$$

Example 1. Let Q be the set of all rational numbers. Then 0 is an identity element for the binary operation of addition on Q since $0 + a = a + 0 = a$ for every $a \in$ Q.

Example 2. For the binary operation of addition on N, there is no identity element, but 1 is an identity element with respect to multiplication. We have $1 \cdot a = a \cdot 1 = a$, for every $a \in$ N.

Inversible elements for a binary operation with identity.

An element a *of a set* A *is said to be inversible for a binary operation* $\circ$ *with identity* e *if* $\exists\ b \in A$ *such that*

$$a \circ b = e = b \circ a.$$

Also then b *is said to be an inverse of* a *and is denoted by* a^{-1}.

The inversible elements in A are also called the *units* in A

The identity element is always inversible and is its own inverse, since $e \circ e = e \circ e = e$. Thus $e^{-1} = e$.

Example 1. The inverse with respect to addition, *or additive inverse* of $x \in$ **R** is $-x$, since $x + (-x) = (-x) + x = 0$, where 0 is the additive identity element of **R**.

Example 2. For the operation of multiplication on Q, 1 is the identity and every element other than 0 is inversible.

Example 3. For the operation of multiplication on N, 1 is the identity and no element except 1 is inversible.

Exercises

1. Define a binary operation. Show that the relation $\circ$ given by $a \circ b = a^b$ is a binary operation on the set of natural numbers. is this binary operation associative ? (Kumayon 1978)

2. Let S be a non-empty set and $\circ$ be a binary operation on S defined by $x \circ y = x$; $x, y \in S$.

Determine whether $\circ$ is commutative and associative.

(Sambalpur 1977)

§ 28. **Relation. Definition.** *Let A and B be two sets. A relation from A to B is a subset of $A \times B$. Symbolically, R is a relation from A to B iff $R \subseteq A \times B$.*

Example 1. Let P be a well-defined set of persons, say, P be the set of all persons living in a certain locality of Meerut. Then the statement "x is a son of y where $x, y \in P$" determines a relation in P. If we denote this relation by R, then R will be the set of all ordered pairs of people belonging to P in which the first co-ordinate is a son of the second co-ordinate. Obviously R will be a subset of $P \times P$. It must be noted that the relation R is the *set R* and not the verbal phrase, 'is a son of'

Several other relations may be defined in the above set P. For example, the statements "x is a daughter of y", "x is the father of y", "x is the mother of y", "x is the husband of y" all determine relations in P.

Example 2. Let **I** be the set of all integers. The statement "x is less than y where $x, y \in \mathbf{I}$" determines a relation in **I**. If we denote this relation by R, then we may describe the set R in the set builder notation as $R=\{(x, y) : x \in \mathbf{I}, y \in \mathbf{I}, x < y\}$.

Suppose R is a relation and (x, y) is an element of the relation R. If $(x, y) \in R$, then sometimes we write xRy which is read as "x is R-related to y". If $(x, y) \notin R$, then sometimes we write $x \not R y$ which is read as "x is not R-related to y". Let **I** be the set of all integers and let $R=\{(x, y) : x \in \mathbf{I}, y \in \mathbf{I}, x < y\}$. Since 2 is less than 3, therefore the ordered pair $(2, 3) \in R$ and we write $2 R 3$. On the other hand 4 is not less than 3 Therefore the ordered pair $(4, 3) \notin R$ and we write $4 \not R 3$.

§ 29. **Domain and Range of a relation.** Let R be a relation from A to B, *i e.*, let R be a subset of $A \times B$. The **domain** D of the relation R is the set of all first elements of the ordered pairs which belong to R. Symbolically,

$$D=\{x : x \in A \text{ and } (x, y) \in R \text{ for some } y \in B\}.$$

The **range** E of the relation R is the set of all second elements f the ordered pairs which belong to R. Symbolically,

$$E=\{y : y \in B \text{ and } (x, y) \in R \text{ for some } x \in A\}.$$

Obviously $D \subseteq A$ and $E \subseteq B$.

Example. Let $A=\{1, 2, 3, 4\}$ and $B=\{a, b, c\}$.

Every subset of $A \times B$ is a relation from A to B. So if $R=\{(2, a), (4, a), (4, c)\}$, then the domain of R is the set $\{2, 4\}$ and the range of R is the set $\{a, c\}$.

§ 30. Total number of distinct relations from a set A to a set B. Suppose the set A has m elements and the set B has n elements. Then the product set $A\times B$ will have mn elements. Therefore the power set of $A\times B$ *i.e.*, $P(A\times B)$ will have 2^{mn} elements. Thus $A\times B$ has 2^{mn} different subsets. Now every subset of $A\times B$ is a relation from A to B. Hence we shall have 2^{mn} different relations from A to B.

§ 31. Inverse Relation . Definition. *Let R be a relation from A to B. The inverse relation of R, denoted by R^{-1}, is a relation from B to A defined by*

$R^{-1}=\{(y, x) : y \in B, x \in A, (x, y) \in R\}$.

In other words, the inverse relation R^{-1} consists of those ordered pairs which when reversed belong to R. Obviously xRy iff $yR^{-1}x$. Thus $(x, y) \in R \Leftrightarrow (y, x) \in R^{-1}$.

Example. Let $A=\{a, b, c\}$, $B=\{1, 2, 3\}$ and

$R=\{(a, 1), (a, 3), (b, 3), (c, 3)\}$. Then

$R^{-1}=\{(1, a), (3, a), (3, b), (3, c)\}$.

Theorem. *If R is a relation from A to B, then the domain of R is identical with the range of R^{-1} and the range of R is identical with the domain of R^{-1}.*

Proof. Let $y \in$ domain of R^{-1}. Then $y \in B$, and $\exists\, x \in A$ such that $(y, x) \in R^{-1}$.

Now $(y, x) \in R^{-1} \Rightarrow (x, y) \in R \Rightarrow y \in$ Range of R.

Thus $y \in$ domain of $R^{-1} \Rightarrow y \in$ Range of R.

$\therefore$ domain of $R^{-1} \subseteq$ range of R.

Similarly range of $R \subseteq$ domain of R^{-1}.

Hence Domain of $R^{-1}=$Range of R.

Similarly we can prove that Domain of $R=$Range of R^{-1}.

Note. If R be a relation from A to B, then it can be easily proved that $(R^{-1})^{-1}=R$.

§ 32. Difference between Relations and Functions.

Suppose A and B are two sets. Let f be a function from A to B. Then from our definition of function, f is a subset of $A\times B$ in which each $a\in A$ appears in one and only one ordered pair belonging to f. In other words f is a subset of $A\times B$ satisfying the following two conditions :

(i) for each $a \in A$, $(a, b) \in f$ for some $b\in B$,

(ii) if $(a, b) \in f$ and $(a, b') \in f$ then $b=b'$.

On the other hand every subset of $A\times B$ is a relation from A to B. *Thus every function is a relation but every relation is not a* *function.* If R is a relation from A to B then domain of R may be a subset of A. But if f is a function from A to B, then domain of f is equal to A. In a relation from A to B an element of A may be related to more than one element in B. Also there may be some elements of A which may not be related to any element in B. But in a function from A to B each element of A must be associated to one and only one element of B.

Example. Let $A=\{1, 2, 3, 4\}$, $B=\{a, b, c\}$.

Let $R=\{(1, a), (2, a), (3, b), (4, b)\}$.

Then R is a function from A to B. Obviously R is also a relation from A to B But consider the subset S of $A \times B$ given by

$$S=\{(1, a), (2, b), (1, c), (3, a), (4, b)\}.$$

Here S is a relation from A to B. But S is not a function from A to B. The obvious reason is that the element $1 \in A$ is associated to two different elements a and $c \in B$.

§ 33. **Relations in a set.** Let R be a relation from A to B. If $B=A$, then we do not speak that R is a relation from A to A, but we say that R is a relation in A. *Thus a relation in a set A is a subset of the Cartesian product $A \times A$.*

Identity relation in a set. **(Gorakhpur 1970)**

Definition. *Let A be a set. The relation I_A defined by $I_A=\{(x, y) : x \in A, y \in A, x=y\}$ is called the identity relation in A.* Thus the identity relation in a set A is the set of the ordered pairs (x, y) of $A \times A$ for which $x=y$. If $A=\{1, 2, 3, 4, 5\}$, then

$$I_A=\{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5)\}.$$

Universal relation in a set. *Let A be any set and R be the set $A \times A$. Then R is called the universal relation in A.*

Void relation in a set. Every subset of $A \times A$ is a relation in A. Since $\emptyset$ is also a subset of $A \times A$, therefore the null set $\emptyset$ is also a relation in A. This relation is called the void relation in A.

§ 34. **Properties of Relations in a Set.** We often come across the following special types of relations in a set :

1. **Reflexive relations.** *Let R be a relation in a set A i.e., let R be a subset of $A \times A$ Then R is called a reflexive relation if*

$$(a, a) \in R, \ \forall \ a \in A.$$

Thus R is reflexive if we have aRa, $\forall \ a \in A$.

A relation R in a set A **is not reflexive** if there is at least one element $a \in A$, such that $(a, a) \notin R$.

Example 1. Let $A=\{1, 2, 3, 4\}$. Then (*i*) The relation $R_1=\{(1, 1), (2, 4), (3, 3), (4, 1), (4, 4)\}$ in A is not reflexive since $2 \in A$ and $(2, 2) \notin R_1$.

(*ii*) The relation $R_2=\{(1, 1), (1, 3), (2, 2), (3, 3), (4, 4), (3, 4)\}$ in A is reflexive since $(a, a) \in R_2$ for every $a \in A$.

Example 2. Let N be the set of all natural numbers. The relation R in N defined by "x is a divisor of y" is reflexive since each natural number is a divisor of itself. It should be noted that it is usual to write $a \mid b$, $a \in$ N, $b \in$ N to denote that a is a divisor of b.

The relation in N defined by "$x < y$" is not reflexive, since a is not less than a for any natural number a.

But the relation in N defined by "$x \leqslant y$" is reflexive since $a \leqslant a$ for every $a \in$ N.

2. **Symmetric Relations.** *Let R be a relation in a set A, i.e., let R be a subset of $A \times A$. Then R is said to be a symmetric relation if $(a, b) \in R \Rightarrow (b, a) \in R$.*

Thus R is symmetric if we have bRa whenever we have aRb.

A relation R in a set A is **not symmetric** if there exist two distinct elements $a, b \in A$, such that aRb, but $b \not R a$.

Example 1. Let L be the set of all straight lines in a plane. The relation R in L defined by "x is perpendicular to y" is symmetric, since if straight line a is perpendicular to straight line b, then b is also perpendicular to a. Thus $aRb \Rightarrow bRa$. The relation in L defined by "x is parallel to y" is also symmetric.

Example 2. Let $A=\{1, 2, 3, 4\}$ and $R=\{(1, 2), (3, 4), (2, 1), (3, 3)\}$. Here we see that $(3, 4 \in R$ but $(4, 3) \notin R$. Therefore R is not symmetric. On the other hand, the relation $R_1=\{(1, 1), (4, 1), (1, 4)\}$ defined in A is symmetric.

3. **Anti-Symmetric relations.** *Let R be a relation in a set A i.e. let R be a subset of $A \times A$. Then R is said to be an anti-symmetric relation if $(a, b) \in R$ and $(b, a) \in R \Rightarrow a=b$.*

Thus R is anti-symmetric if we have never both aRb and bRa except when $a=b$.

A relation R in a set A is **not anti-symmetric** if there exist elements $a \in A$, $b \in A$, $a \neq b$, such that aRb and bRa.

Example. Let N be the set of all natural numbers. Let R be the relation in N defined by "x is a divisor of y". Then R is anti-symmetric since $a \mid b$ and $b \mid a \Rightarrow a=b$. If $a \neq b$, we cannot have both $a \mid b$ and $b \mid a$.

4. Transitive Relations. *Let R be a relation in a set A, i.e., let R be a subset of $A \times A$. Then R is said to be a transitive relation if $(a, b) \in R$ and $(b, c) \in R \Rightarrow (a, c) \in R$.*

A relation R in a set A is **not transitive** if there exist elements a, b and $c \in A$, not necessarily distinct, such that

$$(a, b) \in R,\ (b, c) \in R \text{ but } (a, c) \notin R.$$

Example 1. Let L be the set of all straight lines in a plane and R be the relation in L defined by "x is perpendicular to y". If straight line a is perpendicular to b and b is perpendicular to c, then a is parallel to c, *i.e.*, a is not perpendicular to c. Thus aRb and $bRc \Rightarrow a\not R c$. Hence R is not transitive.

Example 2. Let N be the set of all natural numbers and let R be the relation in N defined by "x is less than y". Since $a<b$ and $b<c \Rightarrow a<c$, therefore R is transitive.

§ 35. Equivalence Relations. Definition.

Let R be a relation in a set A. Then R is an equivalence relation in A iff

(*i*) *R is reflexive, i.e., $\forall\ a \in A$, aRa.*

(*ii*) *R is symmetric, i.e., $aRb \Rightarrow bRa$.*

(*iii*) *R is transitive, i.e., aRb and $bRc \Rightarrow aRc$.*

The equivalence relation is usually denoted by the symbol $\sim$.

Example 1. Let A be the set of all triangles in a plane. Let R be the relation in A defined as xRy iff x is congruent to y, $x \in A$, $y \in A$. Here we observe that

(*i*) xRx, for every $x \in A$, since every triangle is congruent to itself. Thus R is reflexive.

(*ii*) $xRy \Rightarrow yRx$, since if triangle x is congruent to triangle y then y is congruent to x. Thus R is symmetric.

(*iii*) xRy and $yRz \Rightarrow xRz$, since if triangle x is congruent to y, and triangle y is congruent to z, then triangle x is congruent to z. Thus R is transitive.

Since R is reflexive, symmetric and transitive, therefore R is an equivalence relation.

Example 2. Let **I** be the set of all integers. Let us define a relation R in **I**, such that xRy holds iff $x-y$ is divisible by 5, $x \in$ **I**, $y \in$ **I**, *i.e.* let $R=\{(x, y) : x \in \mathbf{I}, y \in \mathbf{I}, x-y \text{ is divisible by } 5\}$.

Here we have

(*i*) For each $x \in$ **I**, $x-x=0$ and 0 is divisible by 5. Thus $\forall\ x \in$ **I**, we have xRx. Therefore R is reflexive.

(*ii*) Suppose xRy; then $x-y$ is divisible by 5 and hence $(y-x)=-(x-y)$ is also divisible by 5. Thus $xRy \Rightarrow yRx$. Therefore R is symmetric.

(*iii*) Suppose xRy and yRz; then $(x-y)$ and $(y-z)$ are both divisible by 5. Hence 5 is also a divisor of $(x-y)+(y-z)$, *i.e.*, 5 is also a divisor of $(x-z)$. Thus xRy and $yRz \Rightarrow xRz$. Therefore R is transitive.

Since R is reflexive, symmetric and transitive, therefore R is an equivalence relation.

Note. An integer a is said to be congruent to another integer b modulo a fixed positive integer m if $a-b$ is divisible by m *i.e.*, if $\exists$ an integer k such that $a-b=mk$. Symbolically, we write

$$a \equiv b \pmod{m}.$$

Thus in the above example, we have proved that congruence modulo 5 is an equivalence relation in the set of all integers.

§ 36. Equivalence Classes or Equivalence Sets.

Let A be a non-empty set and let R be an equivalence relation in A. Further let a be an arbitrary element of A. The elements $x \in A$ satisfying xRa constitute a subset A_a of A, called an equivalence class of a with respect to R. We shall denote this equivalence class by A_a or by $[a]$ or by $\bar{a}$. Thus symbolically,

A_a or $[a]$ or $\bar{a}=\{x : x \in A$ and $(x, a) \in R$ *i.e.*, $xRa\}$.

Example 1. Let A be the set of all triangles in a plane and let R be an equivalence relation in A defined by "x is congruent to y", $x \in A$, $y \in A$. When $a \in A$ we shall mean by the equivalence class $[a]$ the set of all triangles of A congruent to the triangle a. Similarly when $b \in A$ we shall mean by the equivalence class $[b]$ the set of all triangles of A congruent to the triangle b.

Example 2. Let us determine the equivalence classes in the set **I** of all integers with respect to the equivalence relation 'congruence modulo 5'. The integers congruent to 0 modulo 5 form an equivalence class $\mathbf{I}_0$. The elements of this class are the multiples of 5 *i.e.*, integers expressible as $5k$ for some integer k. The integers congruent to 1 modulo 5 form another equivalence class $\mathbf{I}_1$. The elements of this class are those integers which leave remainder 1 when divided by 5 *i.e.*, the integers expressible as $5k+1$ for some integer k. The integers congruent to 2 modulo 5 form a third equivalence class $\mathbf{I}_2$. This class consists of all integers expressible as $5k+2$ for some integer k. The integers congruent to 3 modulo

5 form a fourth equivalence class I_3. This class consists of all integers expressible as $5k+3$ for some integer k. The integers congruent to 4 modulo 5 form a fifth equivalence class I_4. This class consists of all integers which leave remainder 4 when divided by 5 *i.e.*, which are expressible as $5k+4$ for some integer k. But every integer is expressible as $5k$, $5k+1$, $5k+2$, $5k+3$ or $5k+4$ for some integer k. Thus we see that the set of all integers can be divided into the following five equivalence classes with respect to the relation congruence modulo 5 :

(i) $I_0=\{.., -10, -5, 0, 5, 10, ...\}$
(ii) $I_1=\{..., -9, -4, 1, 6, 11, ...\}$
(iii) $I_2=\{..., -8, -3, 2, 7, 12, ...\}$
(iv) $I_3=\{..., -7, -2, 3, 8, 13, ...\}$
(v) $I_4=\{..., -6, -1, 4, 9, 14, ...\}$.

These classes have the following properties :—

(i) The set I is the union of these five non-empty classes.

(ii) Integers in each class have a relation of congruence modulo 5 with one another.

(iii) Integers in different classes do not have a relation of congruence modulo 5 with one another.

(iv) The classes are mutually disjoint *i.e.*, no two of them have any elements in common.

§ 37. Properties of Equivalence classes.

Let A be a non-empty set and let R be an equivalence relation in A. Let a and b be arbitrary elements in A. Then

(i) $a\in[a]$,

(ii) *If* $b\in[a]$, *then* $[b]=[a]$.

(iii) $[a]=[b]$ *iff* $(a, b)\in R$ *i e., iff* aRb.

(iv) *Either* $[a]=[b]$ or $[a]\cap[b]=\varnothing$ *i.e., two equivalence classes are either disjoint or identical.* **[Allahabad 1980]**

Proof.

(i) Since R is reflexive, we have aRa. But $[a]=\{x : x\in A$ and $xRa\}$. Hence $aRa\Rightarrow a\in[a]$.

(ii) We have $b\in[a]\Rightarrow bRa$.

Let x be any arbitrary element of $[b]$. Then $x\in[b] \Rightarrow xRb$. But R is transitive, therefore xRb and $bRa \Rightarrow xRa\Rightarrow x\in[a]$. Thus $x\in[b]\Rightarrow x\in[a]$. Therefore $[b] \subseteq [a]$.

Again let y be any arbitrary element of $[a]$. Then $y\in[a]\Rightarrow yRa$. Since R is symmetric, therefore $bRa\Rightarrow aRb$.

Now $\quad yRa$ and $aRb\Rightarrow yRb$ $\quad$ [$\because$ R is transitive]
$\Rightarrow y\in[b]$.

Thus $y \in [a] \Rightarrow y \in [b]$. Therefore $[a] \subseteq [b]$.

Finally $[a] \subseteq [b]$ and $[b] \subseteq [a] \Rightarrow [a]=[b]$.

(iii) Suppose $[a]=[b]$. Then to prove that aRb.

Since R is reflexive, therefore we have aRa.

$$\begin{aligned} \text{Now} \quad aRa &\Rightarrow a \in [a] \\ &\Rightarrow a \in [b] \qquad [\because [a]=[b]] \\ &\Rightarrow aRb. \end{aligned}$$

Thus $[a]=[b] \Rightarrow aRb$.

Converse. Suppose that aRb. Then to prove that $[a]=[b]$.

Let x be any arbitrary element of $[a]$. Then xRa. But it is given that aRb. Therefore

$$\begin{aligned} xRa \text{ and } aRb &\Rightarrow xRb \qquad [\because R \text{ is transitive}] \\ &\Rightarrow x \in [b]. \end{aligned}$$

Thus $x \in [a] \Rightarrow x \in [b]$. Therefore $[a] \subseteq [b]$.

Again let y be any arbitrary element of $[b]$. Then

$$y \in [b] \Rightarrow yRb.$$

Now we are given that aRb From this we have bRa, since R is symmetric. Now

$$\begin{aligned} yRb \text{ and } bRa &\Rightarrow yRa \qquad [\because R \text{ is transitive}] \\ &\Rightarrow y \in [a]. \end{aligned}$$

Thus $y \in [b] \Rightarrow y \in [a]$. Therefore $[b] \subseteq [a]$.

Hence $[a] \subseteq [b]$ and $[b] \subseteq [a] \Rightarrow [a]=[b]$.

Finally, since $[a]=[b] \Rightarrow aRb$ and $aRb \Rightarrow [a]=[b]$, therefore $[a]=[b]$ if and only if aRb.

(iv) If $[a] \cap [b]=\varnothing$, then we are nothing to prove. So let us suppose that $[a] \cap [b] \neq \varnothing$. Then to prove that $[a]=[b]$.

Since $[a] \cap [b] \neq \varnothing$, therefore there exists an element $x \in A$ such that $x \in [a] \cap [b]$.

$$\begin{aligned} \text{Now} \quad x \in [a] \cap [b] &\Rightarrow x \in [a] \text{ and } x \in [b] \\ &\Rightarrow xRa \text{ and } xRb \\ &\Rightarrow aRx \text{ and } xRb \\ &\qquad [\text{Since } R \text{ is symmetric, therefore } xRa \Rightarrow aRx] \\ &\Rightarrow aRb \qquad [\because R \text{ is transitive}] \\ &\Rightarrow [a]=[b] \qquad [\text{by part (iii)}] \end{aligned}$$

Thus $[a] \cap [b] \neq \varnothing \Rightarrow [a]=[b]$.

In other words if $[a] \neq [b]$, then $[a] \cap [b]=\varnothing$.

§ 38. Partitions. Definition.

Let S be a non-empty set. A set $P=\{A, B, C, \ldots\}$ of non-empty subsets of S will be called a partition of S if *(i)* $A \cup B \cup C \ldots = S$ *i.e., the set* S *is the union of the sets in* P *and*

(ii) the intersection of every pair of distinct subsets of $S \in \mathrm{P}$ *is the null set i.e., if* A *and* $B \in \mathrm{P}$ *then either* $A = B$ *or* $A \cap B = \varnothing$.

Example 1. Consider the set $S = \{1, 2, \ldots 9, 10\}$ and its subsets $B_1 = \{1, 3\}$, $B_2 = \{7, 8, 10\}$, $B_3 = \{2, 5, 6\}$, $B_4 = \{4, 9\}$.

The set $\mathrm{P} = \{B_1, B_2, B_3, B_4\}$ is such that

(i) B_1, B_2, B_3, B_4 are all non-empty subsets of S.

(ii) $B_1 \cup B_2 \cup B_3 \cup B_4 = S$, and

(iii) For any sets B_i, either $B_i = B_j$ or $B_i \cap B_j = \varnothing$.

Hence the set $\{B_1, B_2, B_3, B_4\}$ is a partition of S.

Example 2. Let $\mathbf{I}$ be the set of all integers.

We know that $x \equiv y \pmod 5$ is an equivalence relation in $\mathbf{I}$. Consider the set of five equivalence classes $\mathbf{I}_0, \mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3, \mathbf{I}_4$, where

$$\mathbf{I}_0 = \{\ldots, -10, -5, 0, 5, 10, \ldots\}$$
$$\mathbf{I}_1 = \{\ldots, -9, -4, 1, 6, 11, \ldots\}$$
$$\mathbf{I}_2 = \{\ldots, -8, -3, 2, 7, 12, \ldots\}$$
$$\mathbf{I}_3 = \{\ldots, -7, -2, 3, 8, 13, \ldots\}$$
$$\mathbf{I}_4 = \{\ldots, -6, -1, 4, 9, 14, \ldots\}.$$

We observe that

(i) the sets $\mathbf{I}_0, \mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3$ and $\mathbf{I}_4$ are non-empty,

(ii) the sets $\mathbf{I}_0, \mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3$ and $\mathbf{I}_4$ are pairwise disjoint,

(iii) $\mathbf{I} = \mathbf{I}_0 \cup \mathbf{I}_1 \cup \mathbf{I}_2 \cup \mathbf{I}_3 \cup \mathbf{I}_4$

Hence $\{\mathbf{I}_0, \mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3, \mathbf{I}_4\}$ is a partition of $\mathbf{I}$.

§ 39. **Relation induced by a partition of a set.** Corresponding to any partition of a set S, we can define a relation R in S by the requirement that xRy iff x and y belong to the same subset of S belonging to the partition. The relation R is then said to be induced by the partition.

Example. Consider the subsets

$$A = \{3, 6, 9, \ldots, 24\},\ B = \{1, 4, 7, \ldots, 25\},$$
$$C = \{2, 5, 8, \ldots, 23\} \text{ of } S = \{1, 2, 3, 4, \ldots, 25\}.$$

Obviously $A \cup B \cup C = S$ and $A \cap B = A \cap C = B \cap C = \varnothing$,

so that $\{A, B, C\}$ is a partition of S. If R be the relation induced by this partition, then we have xRy iff x and y belong to the same subset A, B, C.

§ 40. **Fundamental Theorem on Equivalence Relations.** *An equivalence relation R in a non-empty set S determines a partition of S and conversely, a partition of S defines an equivalence relation in S.* **(Sagar 1968 ; Kolhapur 73)**

Proof. Let R be an equivalence relation in S. Let $\mathbf{A}$ be the set of equivalence classes of S with respect to R *i.e.*, let

$$\mathbf{A}=\{[a] : a\in S\}$$

where $$[a]=\{x : x\in S \text{ and } xRa\}.$$

Now R is an equivalence relation. Therefore $\forall\ a \in S$, we have aRa. Hence $a\in[a]$ and thus $[a]\neq\varnothing$.

Further every element a of S is an element of the equivalence class $[a]$ in $\mathbf{A}$ From this we conclude that $S=\bigcup\limits_{a\in S}[a]$.

Finally, if $[a]$ and $[b]$ are two equivalence classes then either

$[a]=[b]$ or $[a]\cap[b]=\varnothing$. [For proof see § 37 part iv].

Hence $\mathbf{A}$ is a partition of S. Thus we see that an equivalence relation R in S decomposes the set S into equivalence classes any two of which are either equal or mutually disjoint.

Converse. Let $P=\{T_a, T_b, T_c,\ldots\}$ be any partition of S. If $p, q \in S$, let us define a relation R in S by pRq iff there is a T_i in the partition such that $p, q \in T_i$.

Now $S=T_a\cup T_b\cup T_c\ldots$. Therefore $\forall\ x \in S$, there exists $T_i \in P$ such that $x \in T_i$.

Hence $x \in T_i$ and $x \in T_i$ means xRx. Thus $\forall\ x \in S$, we have xRx and thus R is reflexive.

Again if we have xRy, then there exists $T_i \in P$ such that $x \in T_i$ and $y \in T_i$.

But $x \in T_i$ and $y \in T_i \Rightarrow y \in T_i$ and $x \in T_i \Rightarrow yRx$.

Therefore R is symmetric.

Finally, suppose xRy and yRz. Then by the definition of R there exist subsets T_j and T_k (not necessarily distinct) such that $x, y \in T_j$ and $y, z \in T_k$. Since $y \in T_j$ and also $y \in T_k$, therefore $T_j\cap T_k\neq\varnothing$. But T_j and T_k belong to a partition of S. Therefore $T_j\cap T_k\neq\varnothing$ implies $T_j=T_k$. Now $T_j=T_k$ implies $x, z \in T_j$ and consequently we have xRz. Thus R is transitive.

Since R is reflexive, symmetric and transitive, therefore R is an equivalence relation.

§ 41. **Quotient Set. Definition.**

Let S be any non-empty set and let R be an equivalence relation defined in S. The set of mutually disjoint equivalence classes in *which S is partitioned relatively to the equivalence relation R, is said to be the Quotient set of S for the equivalence relation R, and is denoted by S/R or by* $\overline{S}$.

The quotient set of I for the equivalence relation congruence modulo 5 is the set

$$\mathbf{I}/R=\{\mathbf{I}_0, \mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3, \mathbf{I}_4\}$$

or

$$\mathbf{I}/R=\{[0], [1], [2], [3], [4]\}.$$

§ **42. Partial Order Relations.** *Definition. A relation R in a set S is called a partial order relation iff it satisfies the following three conditions :*

(*i*) $aRa, \forall a \in S$; (*reflexivity*),

(*ii*) aRb *and* $bRa \Rightarrow a=b$, (*anti-symmetry*),

(*iii*) aRb *and* $bRc \Rightarrow aRc$, (*transitivity*).

Example. In the set N of all natural numbers, the relation R defined by aRb iff a divides b is a partial order relation.

We have, $\forall a \in \mathrm{N}$, a is a divisor of a *i.e.*, aRa. Therefore R is reflexive.

Again, if a is a divisor of b then b cannot be a divisor of a unless $a=b$. Thus aRb and $bRa \Rightarrow a=b$. Therefore R is anti-symmetric.

Finally a is a divisor of b and b is a divisor of c implies a is a divisor of c. Therefore R is transitive.

Since R is reflexive, anti-symmetric and transitive, therefore R is a partial order relation.

Exercises

1. Give an example of a relation which is reflexive but is neither symmetric nor transitive. **(Meerut 1977)**

2. Give an example of a relation which is,

(*i*) reflexive, symmetric but not transitive. **(Kolhapur 1973)**

(*ii*) symmetric and transitive but not reflexive.

3 Is the relation 'is brother of' an equivalence relation on a set of human beings ? Why ?

4. Let I be the set of all integers. Let m be any fixed positive integer. Then an integer a is said to be congruent to another integer b modulo m if $a-b$ is divisible by m *i.e.*, if there exists an integer k such that $a-b=km$. Symbolically, we write

$$a \equiv b \pmod{m}.$$

Show that the relation 'congruent modulo m' is an equivalence relation in the set of integers.

5. Define the terms *'inverse, reflexive, symmetric* and *transitive relations'* and illustrate them by examples. Is function a relation ? When is a relation in a set an equivalence relation ? Prove that any two equivalence classes are either identical or they have null intersection. Hence define a quotient set ?

6. Explain fully what you mean by partitions of a set and equivalence relation in a set by means of examples. Find all the partitions of $\{1, 2, 3\}$.

7. Prove that the relation of similarity in the set of all triangles in a plane is an equivalence relation.

8. Consider the set $N \times N$ the set of ordered pairs of natural numbers. Let R be the relation in $N \times N$ which is defined by

$$(a, b)\ R\ (c, d)$$

if and only if

$$ad = bc.$$

Prove that R is an equivalence relation and therefore induces a partition of $N \times N$.

9. Show that if R and R' are symmetric relations in a set A, then $R \cap R'$ and $R \cup R'$ are also symmetric relations in A.

10. If R and S are two equivalence relations, then check $R \cup S$ for

(i) reflexivity, (ii) transitivity and (iii) symmetry.

11. If R and S are equivalence relations in a set X, prove that $R \cap S$ is an equivalence relation in X.

12. Define a relation and a function and give examples to illustrate the difference between the two. Give an example of a relation which is reflexive and transitive but is not symmetric.

Answers

1. Let $A = \{1, 2, 3\}$. Then the relation $R = \{(1, 1), (2, 3), (2, 2), (3, 3), (1, 2)\}$ is reflexive but is neither symmetric nor transitive.

3. No. Suppose x is a female human being. Then x is not a brother of x and so $x \not R x$. Thus the relation is not reflexive. Note that this relation is also not symmetric.

6. Partitions of $\{1, 2, 3\}$ are

(i) $\{\{1, 2, 3\}\}$, (ii) $\{\{1\}, \{2\}, \{3\}\}$, (iii) $\{\{1, 2\}, \{3\}\}$, (iv) $\{\{2, 3\}, \{1\}\}$, (v) $\{\{3, 1\}, \{2\}\}$.

10. $R \cup S$ is reflexive and symmetric but is not necessarily transitive.

2

Groups

§ 1. Binary operation on a set Let G be a non-empty set. Then $G \times G = \{(a, b) : a \in G, b \in G\}$ *If* $f : G \times G \to G$, *then* f *is said to be a binary operation on the set* G. The image of the ordered pair (a, b) under the function f is denoted by $a f b$. Often we use symbols $+$, $\times$, $.$, o, $*$ etc. to denote binary operations on a set. Thus '$+$' will be a binary operation on G iff

$$a+b \in G \ \forall \ a, b \in G \text{ and } a+b \text{ is unique.}$$

Similarly '$*$' will be a binary operation on G iff

$$a * b \in G \ \forall \ a, b \in G \text{ and } a * b \text{ is unique.}$$

A **binary operation** on a set G is sometimes also called **binary composition** in the set G. If '$*$' is a binary composition in G, then $a * b \in G \ \forall \ a, b \in G$. Therefore G is closed with respect to the composition denoted by $*$.

If there is a binary composition in a set G, the most convenient notation to denote this composition is the multiplicative notation. In this notation if $a, b \in G$, then $a \cdot b$ represents the element obtained on multiplying a and b. We shall often omit the dot $\cdot$ placed between a and b. In other words in multiplicative notation we shall simply write ab in place of $a \cdot b$ Thus $ab \in G \ \forall \ a, b \in G$ if the binary composition in G has been denoted multiplicatively.

Example. Addition is a binary operation on the set N of natural numbers. The sum of two natural numbers is also a natural number. Therefore N is closed with respect to addition

i.e., $$a+b \in \mathbf{N} \ \forall \ a, b \in \mathbf{N}.$$

Subtraction is not a binary operation on N. We have $4-7=-3 \notin \mathbf{N}$ whereas $4 \in \mathbf{N}, 7 \in \mathbf{N}$. Thus N is not closed with respect to subtraction. But subtraction is a binary operation on the set of integers I. We have $a-b \in \mathbf{I} \ \forall \ a, b \in \mathbf{I}$.

Note. In the theory of groups we shall be concerned only with binary operations. Therefore we shall often omit the word binary and we shall simply use the word '**operation**'.

§ 2. **Algebraic structure. Definition.** *A non-empty set G equipped with one or more binary operations is called an algebraic structure.* Suppose $*$ is a binary operation on G. Then $(G, *)$ is an algebraic structure. $(\mathbf{N}, +)$, $(\mathbf{I}, +)$, $(\mathbf{I}, -)$, $(\mathbf{R}, +, \cdot)$ are all algebraic structures. Obviously addition and multiplication are both binary operations on the set **R** of real numbers. Therefore $(\mathbf{R}, +, \cdot)$ is an algebraic structure equipped with two operations.

Semi-Group. Definition. *An algebraic structure $(G, *)$ is called a semi-group if the binary operation $*$ is associative in G i.e., if $(a * b) * c = a * (b * c) \; \forall \; a, b, c \in G$.* **(Rajasthan 1976; Kerala 70)**

For example, the set N of all natural numbers is a semi-group with respect to the operation of addition of natural numbers. Obviously addition is an associative operation on N.

Similarly the algebraic structures $(\mathbf{N}, \cdot)$, $(\mathbf{I}, +)$ and $(\mathbf{R}, +)$ are also semi-groups.

§ 3. **Group Definition.**

(I.A.S. 1969; Vikram 76; Garhwal 76; Poona 73; Kolhapur 73; Madras 74; Sagar 77; Guru Nanak 82; Meerut 81)

Let G be a non-empty set equipped with a binary operation denoted by $\cdot$ i.e., $a \cdot b$ or more conveniently ab represents the element of G obtained by applying the said binary operation between the elements a and b of G taken in that order. Then this algebraic structure $(G, \cdot)$ is a group, if the binary operation $\cdot$ satisfies the following postulates :

1. **Closure property** *i.e.,* $ab \in G \; \forall \; a, b \in G$.
2. **Associativity** *i.e.,* $(ab)\, c = a\, (bc) \; \forall \; a, b, c \in G$.
3. **Existence of Identity.** *There exists an element $e \in G$ such that $ea = a = ae \; \forall \; a \in G$. The element e is called the identity.*
4. **Existence of inverse.** *Each element of G possesses inverse. In other words $a \in G \Rightarrow$ there exists an element $b \in G$ such that $ba = e = ab$. The element b is then called the* inverse *of a and we write $b = a^{-1}$. Thus a^{-1} is an element of G such that $a^{-1}a = e = aa^{-1}$.*

Abelian Group or Commutative Group. Definition.

(Madras 1977)

A group G is said to be abelian or commutative if in addition to the above four postulates the following postulate is also satisfied.

5. **Commutativity** *i.e.,* $ab = ba \; \forall \; a, b \in G$.

Note 1. According to our definition of a binary operation if '$\cdot$' is a binary operation on G, we must have $a \cdot b \in G$, $\forall \; a$,

$\in G$. Therefore in our definition of a group there is no necessity of mentioning the closure axiom. It is superfluous there. We have mentioned it there to simply lay emphasis upon the fact that while showing the group postulates in a problem the students should not forget to show the closure axiom.

Note 2. A group is not simply a set but it is an algebraic structure *i.e.*, a set equipped with a binary composition provided the composition satisfies certain postulates. If a group consists of a non-empty set G and a binary composition in G, then we shall often use the same symbol G to denote the group and the underlined set.

Note 3. In our definition of a group we have denoted the composition in G by multiplicative notation. However we can use any symbol like $*$, $\circ$, $+$etc. to denote the composition. If we use the additive notation '$+$' to denote the composition in G, then the inverse of an element $a \in G$ is denoted by the symbol $-a$ *i.e.*, we have $(-a)+a=e=a+(-a)$.

Note 4. If we use multiplicative notation to denote the composition in G, then often we denote the identity by the symbol '1'. Thus 1 is an element of G such that

$$1\,a=a=a1 \quad \forall \quad a \in G.$$

Also in multiplicative notation we often denote the inverse of a by $1/a$. Thus $1/a$ is an element of G such that $\frac{1}{a}\,a=1=a\,\frac{1}{a}$.

In additive notation, we often denote the identity by the symbol '0'. Then 0 is an element of G such that $0+a=a=a+0$

Note 5. In additive notation the element $a+(-b) \in G$ is denoted by $a-b$. In multiplicative notation the element $ab^{-1} \in G$ is denoted by a/b.

§ 4. Finite and infinite groups. Order of a finite group.

(Rajasthan 1977 ; Kanpur 80)

If in a group G the underlying set G consists of a finite number of distinct elements then the group is called a **finite group**, otherwise an **infinite group**. *The number of elements in a finite group is called the* **order** *of the group.* An infinite group is said to be of infinite order.

We shall denote the order of a group G by the symbol $o\,(G)$.

It should be noted that the smallest group for a given composition is the set $\{e\}$ consisting of the identity element e alone.

Some Illustrative Examples.

Example 1. *Show that the set* **I** *of all integers*

$$\ldots, -4, -3, -2, -1, 0, 1, 2, 3, 4, \ldots$$

is a group with respect to the operation of addition of integers.

(Meerut 1970 ; Lucknow 80)

Solution. Closure property. We know that the sum of two integers is also an integer *i.e.*, $a+b \in \text{I} \ \forall \ a, b \in \text{I}$. Thus I is **closed** with respect to addition.

Associativity. We know that addition of integers is an associative composition. Therefore,

$$a+(b+c)=(a+b)+c \ \forall \ a, b, c \in \text{I}.$$

Existence of Identity. The number $0 \in \text{I}$. Also we have $0+a=a=a+0 \ \forall \ a \in \text{I}$. Therefore the integer 0 is the identity.

Existence of Inverse. If $a \in \text{I}$, then $-a \in \text{I}$ Also we have $(-a)+a=0=a+(-a)$. Thus every integer possesses additive inverse.

Therefore I is a group with respect to addition. Since addition of integers is a commutative composition, therefore (I, +) is an abelian group. Also I contains an infinite number of elements. Therefore (I, +) is an abelian group of infinite order.

Example 2. *Show that the set* N *of all natural numbers*

$$1, 2, 3, 4, 5, \ldots$$

is not a group with respect to addition.

Solution. Addition is obviously a binary composition in N *i.e.*, N is closed with respect to addition. Also addition of natural numbers is an associative composition. But there exists no natural number $e \in \text{N}$ such that $e+a=a=a+e \ \forall \ a \in \text{N}$. For the addition of numbers, the number 0 is the identity and $0 \notin \text{N}$. Therefore (N, +) is not a group.

Example 3. *Show that the set*

$$G=\{\ldots, -4m, -3m, -2m, -m, 0, m, 2m, 3m, 4m, \ldots\}$$

of multiples of integers by a fixed integer m is a group with respect to addition. **(Lucknow 1967)**

Solution. Closure property. Let a, b be any two elements of G. Then $a=rm$ and $b=sm$ where r and s are some integers.

Now $a+b=rm+sm=(r+s)\,m$. Since $r+s$ is also an integer, therefore $(r+s)\,m$ *i.e.*, $a+b \in G$. Thus $a+b \in G \ \forall \ a, b \in G$. Therefore G is closed with respect to addition.

Associativity. The elements of G are all integers and we know that the addition of integers is an associative composition.

Existence of Identity. $0 \in G$ and we have $0+a=a=a+0 \ \forall \ a \in G$. Therefore 0 is the identity.

Existence of Inverse. Let rm be an arbitrary element of G where r is some integer.

Then $(-r)\, m \in G$. [$\because$ $-r$ is also an integer]

Also $(-r)\, m+rm=(-r+r)\, m=0m=0$

and $rm+(-r)\, m=(r-r)\, m=0$.

$\therefore$ $(-r)\, m$ is the additive inverse of rm.

Thus every element of G possesses additive inverse.

Hence G is a group with respect to addition.

Example 4. *Prove that the set* Q_0 *of all non-zero rational numbers forms a group under the operation of multiplication of rational numbers.* **(Bhagalpur 1971 ; Meerut 74)**

Solution. Closure property. We know that the product of two non-zero rational numbers is also a non-zero rational number. Therefore Q_0 is closed with respect to multiplication.

Associativity. We know that multiplication of rational numbers is an associative composition.

Existence of Identity. The rational number $1 \in Q_0$. Also we have $1a=a=a1 \ \forall \ a \in Q_0$.

$\therefore$ The rational number 1 is the multiplicative identity.

Existence of Inverse. If $a \in Q_0$, then obviously $1/a \in Q_0$. Also $(1/a)\, a=1=a\, (1/a)$. Therefore $1/a$ is the multiplicative inverse of a.

Hence Q_0 is a group with respect to multiplication.

Since $ab=ba \ \forall \ a, b \in Q_0$, therefore the group is abelian.

Note. The set Q of all rational numbers is not a group with respect to multiplication. The rational number $0 \in Q$, but $\exists$ no rational number $a \in Q$ such that $0a=1$. We know that $0a=0 \ \forall \ a \in Q$. Thus the rational number 0 does not possess multiplicative inverse.

Example 5. *Is the set of all rational numbers* x *such that* $0 < x \leqslant 1$, *a group with respect to ordinary multiplication ?*

(Madras 1975)

Solution. Let $G=\{x : x \text{ is a rational number and } 0 < x \leqslant 1\}$. Then G is not a group with respect to multiplication. If $a \in G$ and $0 < a < 1$, then a does not possess a multiplicative inverse which is an element of G.

Example 6. *Show that the set of all positive rational numbers forms an abelian group under the composition defined by*

$a * b=(ab)/2$. **(Delhi 1970 ; Meerut 71)**

Solution. Let Q_+ denote the set of all positive rational numbers. We define an operation $*$ on Q_+ as follows :

$$a*b=(ab)/2 \ \forall \ a, b \in Q_+.$$

To show that $(Q_+, *)$ is a group.

Closure Property. Since for every $a, b \in Q_+$, $(ab)/2$ is also in Q_+, therefore Q_+ is closed with respect to the operation $*$.

Associativity. Let $a, b, c \in Q_+$. Then

$$(a*b)*c=\left(\frac{ab}{2}\right)*c=\left(\frac{ab}{2}\right)\frac{c}{2}=\frac{a}{2}\left(\frac{bc}{2}\right)=a*\left(\frac{bc}{2}\right)=a*(b*c).$$

Commutativity. Let $a, b \in Q_+$. Then

$$a*b=(ab)/2=(ba)/2=b*a.$$

Existence of Identity. The number e will be the identity element if $e \in Q_+$ and if $e*a=a=a*e \ \forall a \in Q_+$.

Now $e*a=a \Rightarrow (ea)/2=a \Rightarrow (a/2)(e-2)=0$

$\Rightarrow e=2$, since $a \in Q_+ \Rightarrow a \neq 0$.

Now $2 \in Q_+$ and we have $2*a=(2a)/2=a=a*2 \ \forall a \in Q_+$.

$\therefore$ 2 is the identity element.

Existence of Inverse. Let a be any element of Q_+. If the number b is to be the inverse of a, then we must have

$$b*a=e=2 \Rightarrow (ba)/2=2 \Rightarrow b=4/a.$$

Now $a \in Q_+ \Rightarrow 4/a \in Q_+$.

We have $(4/a)*a=4a/2a=2=a*(4/a)$.

Therefore $4/a$ is the inverse of a. Thus each element of Q_+ is inversible.

Hence $(Q_+, *)$ is an abelian group.

Important Note. In the few examples of groups we have just given, the elements of the underlying sets were all numbers. But it should not be confused that we can have groups of numbers only. In our definition of a group, we have not imposed any restriction on the elements of the set G. The elements of G can be numbers, chairs, students, countries or anything. The set G will be a group if it is equipped with a binary composition satisfying certain postulates.

Exercises

1. (i) State the axioms which a set must obey so that it may form a group. **(Meerut 1977, 79)**

   (ii) Show that every group of order 1 is abelian and every group of order 2 is also abelian.

2. Does the set of all odd integers form a group with respect to addition ? **(Meerut 1979)**

   Ans. No.

3. Show that the following are groups :
   (i) the set C of all complex numbers with respect to the operation of addition of complex numbers. **(Rajasthan 1974)**
   (ii) the set of all rational numbers with respect to addition. **(Bhagalpur 1971, Meerut 79)**
   (iii) the set of all real numbers with respect to addition.
   (iv) the set R_0 of all non-zero real numbers with respect to multiplication.
   (v) the set C_0 of all non-zero complex numbers with respect to multiplication. **(Meerut 1979)**
4. Define the order of a group. Show that the set of all even integers with zero is an abelian group with respect to addition.
5. Show that the set of natural numbers does not form a group with addition or multiplication but it forms a semi-group with respect to addition as well as multiplication.
6. Is the set I of integers

   $$\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$$

   a group (i) with respect to subtraction (ii) with respect to multiplication ? **Ans.** (i) No. (ii) No.
7. Show that the set M of complex numbers z with the condition $|z|=1$ forms a group with respect to the operation of multiplication of complex numbers.
8. Show that the set V of all vectors (defined as directed line segments) forms an infinite abelian group with vector addition as composition.
9. Show that the set of vectors defined as directed line segments does not form a group (i) with respect to scalar (dot, product, (ii) with respect to vector (cross) product.
10. Is the set of all even natural numbers a group
    (i) under addition (ii) under multiplication ?
    Ans. (i) No; (ii) No.
11. Let Q_+ be the set of all positive rational numbers and $*$ a binary operation on Q_+ defined by $a * b=\frac{ab}{3}$. Determine the identity element in Q_+ and determine the inverse of a. **(Meerut 1977, 79, 82)**
    Ans. Identity element is 3, and the inverse of a is $9/a$.

12. Show that the set of all positive rational numbers forms an abelian group under the composition defined by

$$a * b = (ab)/4. \qquad \text{(Meerut 1979)}$$

13. Let **R** be the set of all real numbers and $*$ a binary operation on **R** defined by $a * b = a + b + ab$. Determine the identity element in **R** and determine the inverse of a. (Meerut 1976)

Ans. Identity element is 0; if $a \neq -1$, then $a^{-1} = -\dfrac{a}{a+1}$.

14. Show that the set of all rational numbers of the form $2^a\, 3^b$ (a, b integers) is a group with respect to multiplication of rationals.

15. Do the positive irrationals form a group with respect to multiplication? **Ans.** No.

16. Prove that the set $G = \{(\cos \theta + i \sin \theta) : \theta$ runs over all rational numbers$\}$ forms an infinite abelian group with respect to ordinary multiplication.

§ 5. Some General Properties of Groups.

Suppose our group consists of a non-empty set G equipped with a binary operation denoted multiplicatively.

Theorem 1. Uniqueness of identity. *The identity element in a group is unique.*

(Madras 1974, Andhra 77, Lucknow 80, Sagar 77)

Proof. Suppose e and e' are two identity elements of a group G. We have

$$ee' = e \quad \text{if } e' \text{ is identity}$$

and

$$ee' = e' \quad \text{if } e \text{ is identity.}$$

But ee' is a unique element of G.

Therefore $ee' = e$ and $ee' = e' \Rightarrow e = e'$.

Hence the identity element is unique.

Theorem 2. Uniqueness of inverse. *The inverse of each element of a group is unique.*

(Madras 1974, Lucknow 80, Meerut 80, Andhra 77, Sagar 77)

Proof. Let a be any element of a group G and let e be the identity element. Suppose b and c are two inverses of a *i.e.*,

$$ba = e = ab \quad \text{and} \quad ca = e = ac.$$

We have $b(ac) = be$ $\quad [\because\ ac = e]$

$= b$ $\quad [\because\ e \text{ is identity}]$

Also $(ba)c = ec$ $\quad [\because\ ba = e]$

$= c.$ $\quad [\because\ e \text{ is identity}]$

But in a group composition is associative. Therefore $b(ac) = (ba)c$. Hence $b = c$.

Note. The identity element is its own inverse. Since $ee=e$, therefore $e^{-1}=e$.

Theorem 3. *If the inverse of a is a^{-1}, then the inverse of a^{-1} is a i.e.,* $(a^{-1})^{-1}=a$. **(Lucknow 1967)**

Proof. If e is the identity element, we have

$a^{-1}a=e$ [by definition of inverse]

$\Rightarrow (a^{-1})^{-1}[a^{-1}a]=(a^{-1})^{-1}e$ [multiplying both sides on the left by $(a^{-1})^{-1}$ which is necessarily an element of G because a^{-1} is an element of G]

$\Rightarrow [(a^{-1})^{-1}a^{-1}]a=(a^{-1})^{-1}$ [$\because$ composition in G is associative and e is identity element]

$\Rightarrow ea=(a^{-1})^{-1}$ [$\because$ $(a^{-1})^{-1}$ is inverse of a^{-1}]

$\Rightarrow a=(a^{-1})^{-1}\Rightarrow(a^{-1})^{-1}=a.$

Note. If we had used additive notation to denote the composition in G, the statement of this result would have been

$$-(-a)=a.$$

Theorem 4. *To prove that* $(ab)^{-1}=b^{-1}a^{-1}\ \forall\ a, b\in G$ *i.e., the inverse of the product of two elements of a group G is the product of the inverses taken in the reverse order.*

(Meerut 1975, Vikram 78, Garhwal 76, Kolhapur 73, Berhampur 77, Madras 77)

Proof. Suppose a and b are elements of G. If a^{-1} and b^{-1} are inverses of a and b respectively, then

$a^{-1}a=e=aa^{-1}$ where e is the identity element

and $b^{-1}b=e=bb^{-1}$.

Now $(ab)(b^{-1}a^{-1})=[(ab)b^{-1}]a^{-1}$ [$\because$ composition is associative]

$=[a(bb^{-1})]a^{-1}$ [by associativity]

$=(ae)a^{-1}$ [$\because$ $bb^{-1}=e$]

$=aa^{-1}$ [$\because$ $ae=a$]

$=e$ [$\because$ $aa^{-1}=e$]

Also $(b^{-1}a^{-1})(ab)=b^{-1}[a^{-1}(ab)]$ [by associativity]

$=b^{-1}[(a^{-1}a)b]=b^{-1}(eb)=b^{-1}b=e.$

Thus we have $(b^{-1}a^{-1})(ab)=e=(ab)(b^{-1}a^{-1})$.

$\therefore$ by definition of inverse, we have $(ab)^{-1}=b^{-1}a^{-1}$.

If the group is commutative, then we shall have

$(ab)^{-1}=a^{-1}b^{-1}$, since $b^{-1}a^{-1}=a^{-1}b^{-1}$.

Note 1. In additive notation the statement of this theorem will be $-(a+b)=(-b)+(-a)$.

Note 2. The rule given in this theorem is known as the *reversal rule*. It can be generalised by induction as follows :

If $a, b, c, \ldots, k$ are in G, then

$$(abc\ldots jk)^{-1}=k^{-1}j^{-1}\ldots c^{-1}b^{-1}a^{-1}.$$

Theorem 5. Cancellation laws hold good in a group. *If a, b, c are any elements of G, then*

$$ab=ac \Rightarrow b=c \qquad \textbf{(Left cancellation law)}$$

and $$ba=ca \Rightarrow b=c. \qquad \textbf{(Right cancellation law)}$$

(Allahabad 1970; Berhampur 77)

Proof. $a \in G \Rightarrow \exists\, a^{-1} \in G$ such that $a^{-1}a=e=aa^{-1}$ where e is the identity element.

Now $ab=ac \Rightarrow a^{-1}(ab)=a^{-1}(ac)$ [multiplying both sides on the left by a^{-1}]

$\Rightarrow (a^{-1}a)\,b=(a^{-1}a)\,c$ [by associativity]

$\Rightarrow eb=ec$ [$\because$ $a^{-1}a=e$]

$\Rightarrow b=c$ [$\because$ e is identity]

Also $ba=ca \Rightarrow (ba)\,a^{-1}=(ca)\,a^{-1}$

$\Rightarrow b\,(aa^{-1})=c\,(aa^{-1}) \Rightarrow be=ce \Rightarrow b=c.$

Note. In additive notation these results can be written as $a+b=a+c \Rightarrow b=c$ and $b+a=c+a \Rightarrow b=c$.

Theorem 6. *If a, b are any two elements of a group G, then the equations $ax=b$ and $ya=b$ have unique solutions in G.*

(Nagpur 1975; Madras 77)

Proof. $a \in G \Rightarrow \exists\, a^{-1} \in G$ such that $a^{-1}a=e=aa^{-1}$ where e is the identity element.

$\therefore$ $a \in G, b \in G \Rightarrow a^{-1} \in G, b \in G$

$\Rightarrow a^{-1}b \in G.$ [by closure property]

Now substituting $a^{-1}b$ for x in the left hand side of the equation $ax=b$, we have

$$a\,(a^{-1}b)=(aa^{-1})\,b=eb=b.$$

Thus $x=a^{-1}b$ is a solution in G of the equation $ax=b$.

To show that the solution is unique, let us suppose that $x=x_1$ and $x=x_2$ are two solutions of the equation $ax=b$. Then $ax_1=b$ and $ax_2=b$. Therefore $ax_1=ax_2$. By left cancellation law this gives $x_1=x_2$. Therefore the solution is unique.

Now to prove that the equation $ya=b$ has a unique solution in G. We have $a \in G, b \in G \Rightarrow ba^{-1} \in G$.

Now $(ba^{-1})\,a=b\,(a^{-1}a)=be=b$.

$\therefore$ $y=ba^{-1}$ is a solution in G of the equation $ya=b$.

Suppose y_1 and y_2 are two solutions of this equation. Then

$y_1a=b$ and $y_2a=b$. Therefore $y_1a=y_2a$. By right cancellation law this gives $y_1=y_2$. Therefore the solution is unique.

§ 6. **Definition of a group based upon Left Axioms.** *Let G be a non-empty set equipped with a binary operation denoted by • i.e., a•b or more conveniently ab represents the element of G obtained by applying the said binary operation between the elements a and b of G taken in that order. Then this algebraic structure (G, •) is a group if the binary operation • satisfies the following postulates.*

1. **Closure property** *i.e.,* $ab \in G \ \forall \ a, b \in G$.
2. **Associativity** *i.e.,* $(ab)\, c=a\,(bc) \ \forall \ a, b, c \in G$.
3. **Existence of Left Identity.** *There exists an element* $e \in G$ *such that* $ea=a \ \forall \ a \in G$. *The element e is called the* **left identity.**
4. **Existence of Left Inverse.** *Each element of G possesses left inverse. In other words* $a \in G \Rightarrow$ *there exists an element* $a^{-1} \in G$ *such that* $a^{-1}\, a=e$. *The element* a^{-1} *is the left inverse of a.*

We shall now prove that this definition of a group and the classical definition of a group given in § 3 are equivalent. Obviously if the postulates of a group given in § 3 hold good, the postulates of a group given in this definition will also hold good.

If the postulates of a gruop given in this definition hold good, then the postulates given in § 3 will also hold good if starting with left axioms we prove that the left identity is also the right identity and the left inverse of an element is also the right inverse. First we shall prove the existence of left cancellation law and then we shall prove the other two results.

Theorem 1. **(Left Cancellation Law).** *If a, b, c are in G, then*

$$ab=ac \Rightarrow b=c.$$

Proof. Since $a \in G$, therefore $\exists\ a^{-1} \in G$ such that $a^{-1}\, a=e$ where e is the left identity.

Now $\quad ab=ac \Rightarrow a^{-1}\,(ab)=a^{-1}\,(ac)$

$$\Rightarrow (a^{-1}\, a)\, b=(a^{-1}\, a)\, c \qquad \text{[by associativity]}$$

$$\Rightarrow eb=ec \qquad [\because\ a^{-1} \text{ is left inverse of } a]$$

$$\Rightarrow b=c \qquad [\because\ e \text{ is left identity}]$$

Theorem 2. *The left identity is also the right identity i.e., if e is the left identity, then also* $ae=a \ \forall \ a \in G$. **(Meerut 1968)**

Proof. Let $a \in G$ and e be the left identity. Since a possesses left inverse, therefore there exists $a^{-1} \in G$ such that $a^{-1}\, a=e$.

We have $a^{-1}\,(ae)=(a^{-1}\, a)\, e \qquad$ [by associativity]

$$=ee \qquad [\because\ a^{-1}\, a=e]$$

$$=e \qquad [\because\ e \text{ is left identity}]$$

$=a^{-1}a.$ $[\because a^{-1}a=e]$

Now $a^{-1}(ae)=a^{-1}a \Rightarrow ae=a$, by left cancellation law.

$\therefore$ e is also the right identity.

$\therefore$ e is also the identity *i.e.*, $ea=a=ae \ \forall \ a \in G$.

Theorem 3. *The left inverse of an element is also its right inverse i.e., if* a^{-1} *is the left inverse of* a, *then also* $aa^{-1}=e$.

(Kanpur 1980; Meerut 74)

Proof. Let $a \in G$ and e be the identity element. Let a^{-1} be the left inverse of a *i.e.*, $a^{-1}a=e$. To prove that $aa^{-1}=e$.

We have $a^{-1}(aa^{-1})=(a^{-1}a)a^{-1}$ [by associativity]

$=ea^{-1}$ $[\because a^{-1}a=e]$

$=a^{-1}$ $[\because e$ is left identity$]$

$=a^{-1}e.$ $[\because e$ is also right identity$]$

Now $a^{-1}(aa^{-1})=a^{-1}e$

$\Rightarrow aa^{-1}=e.$ [by left cancellation law]

$\therefore$ a^{-1} is also the right inverse of a. Hence a^{-1} is the inverse of a, *i.e.*, $a^{-1}a=e=aa^{-1}$.

Note 1. In order to prove that a non-empty set G equipped with a binary operation is a group it is sufficient to prove that the operation is associative, the left identity exists and the left inverse of each element of G exists

Note 2. We can also define a group with the help of right axioms only. However we cannot assume the existence of left identity and the existence of right inverse or we cannot assume the existence of right identity and the existence of left inverse.

Some More Examples on Groups

Example 1. *Show that the set* $G=\{a+b\sqrt{2} : a, b \in \mathbb{Q}\}$ *is a group with respect to addition.* **(Meerut 1979)**

Solution. Closure Property. Let x, y be any two elements of G. Then $x=a+b\sqrt{2}$, $y=c+d\sqrt{2}$ where $a, b, c, d \in \mathbb{Q}$.

Now $x+y=(a+b\sqrt{2})+(c+d\sqrt{2})=(a+c)+(b+d)\sqrt{2}$. Since $a+c$ and $b+d$ are elements of $\mathbb{Q}$, therefore $(a+c)+(b+d)\sqrt{2} \in G$. Thus $x+y \in G \ \forall \ x, y \in G$. Therefore G is closed with respect to addition.

Associativity. The elements of G are all real numbers and the addition of real numbers is associative.

Existence of left identity. We have $0+0\sqrt{2} \in G$ since $0 \in \mathbb{Q}$. If $a+b\sqrt{2}$ is any element of G, then

$(0+0\sqrt{2})+(a+b\sqrt{2})=(0+a)+(0+b)\sqrt{2}=a+b\sqrt{2}.$

$\therefore$ $0+0\sqrt{2}$ is the left identity.

Existence of left Inverse. We have

$a+b\sqrt{2} \in G \Rightarrow (-a)+(-b)\sqrt{2} \in G$ since $a, b \in \mathbf{Q} \Rightarrow -a, -b \in \mathbf{Q}$.

Now $[(-a)+(-b)\sqrt{2}]+[a+b\sqrt{2}]$
$=[(-a)+a]+[(-b)+b]\sqrt{2}=0+0\sqrt{2}=$ the left identity.

$\therefore$ $(-a)+(-b)\sqrt{2}$ is the left inverse of $a+b\sqrt{2}$.

Hence G is a group with respect to addition.

Example 2. *Do the following sets form groups with respect to the binary operation $*$ defined on them as follows :*

(*i*) *the set* **I** *of all integers with operation defined by*

$$a * b=a+b+1. \qquad \textbf{(Madras 1975)}$$

(*ii*) *the set* $\mathbf{Q}_1$ *of all rational numbers other than* 1, *with the operation defined by* $a * b=a+b-ab$. **(Kanpur 69)**

(*iii*) *the set* $\mathbf{Q}'$ *of all rational numbers other than* -1 *with the operation defined by* $a * b=a+b+ab$. **(I. A. S. 71; Poona 73)**

Solution.

(i) **Closure Property.** We have $a \in \mathbf{I}, b \in \mathbf{I} \Rightarrow a+b+1$ *i.e.*, $a * b \in \mathbf{I}$ Therefore **I** is closed with respect to the operation $*$.

Associativity. If $a, b, c \in \mathbf{I}$, then

$(a * b) * c=(a+b+1) * c=(a+b+1)+c+1=a+b+c+2.$

Also $a * (b * c)=a * (b+c+1)=a+(b+c+1)+1=a+b+c+2.$

$\therefore$ $(a * b) * c=a * (b * c) \;\forall\; a, b, c \in \mathbf{I}$.

Existence of left identity. $e \in \mathbf{I}$ will be the left identity if $e * a=a \;\forall\; a \in \mathbf{I}$.

Now $e * a=e+a+1.$

$\therefore$ $e+a+1=a \Rightarrow e=-1.$

Since $-1 \in \mathbf{I}$ and we have for any $a \in \mathbf{I}$

$$(-1) * a=-1+a+1=a,$$

therefore -1 is the left identity element.

Existence of left Inverse. If $a \in \mathbf{I}$, then $b \in \mathbf{I}$ will be left inverse of a if $b * a=-1$ (the left identity).

Now $b * a=-1 \Rightarrow b+a+1=-1 \Rightarrow b=-2-a.$

Now $a \in \mathbf{I} \Rightarrow -2-a \in \mathbf{I}$.

Also $(-2-a) * a=(-2-a)+a+1=-1.$

$\therefore$ $-2-a$ is the left inverse of a.

Also $a * b=a+b+1=b+a+1=b * a$. Therefore the composition is also commutative.

Hence **I** is an infinite abelian group for the given composition.

(ii) **Closure Property.** Let $a, b \in \mathbf{Q}_1$. Then a and b are rational numbers such that $a \neq 1$, $b \neq 1$.

Now $a * b=a+b-ab$ which is also a rational number and it cannot be equal to 1,

because $a+b-ab=1 \Rightarrow a+b-ab-1=0 \Rightarrow (a-1)(1-b)=0$
$\Rightarrow a=1$ or $b=1$ which is not so.

$\therefore\ a * b \in Q_1\ \forall\ a, b \in Q_1$. Hence Q_1 is closed with respect to the given composition.

Associativity. If $a, b, c \in Q_1$, then

$$(a * b) * c=(a+b-ab) * c=(a+b-ab)+c-(a+b-ab)\,c$$
$$=a+b+c-ab-ac-bc+abc.$$

Also $a * (b * c)=a * (b+c-bc)=a+(b+c-bc)-a\,(b+c-bc)$
$$=a+b+c-ab-ac-bc+abc.$$

$\therefore\ a * (b * c)=(a * b) * c\ \forall\ a, b, c \in Q_1$.

Existence of left Identity. Let $e \in Q_1$ *i.e.*, let e be a rational number and $e \neq 1$. Then e will be the left identity iff $\forall a \in Q_1$ we have $e * a=a \Leftrightarrow e+a-ea=a \Leftrightarrow e-ea=0 \Leftrightarrow e(1-a)=0 \Leftrightarrow e=0$. [Note that $a \in Q_1 \Rightarrow a \neq 1$]. Now $0 \in Q_1$; so 0 is the left identity.

Existence of left Inverse. Let $a \in Q_1$ *i.e.*, let a be a rational number and $a \neq 1$. Now $b \in Q_1$ will be the left inverse of a iff

$b * a=0 \Leftrightarrow b+a-ba=0 \Leftrightarrow b\,(a-1)=a \Leftrightarrow b=\dfrac{a}{a-1}$, since $a \neq 1$.

Now $\dfrac{a}{a-1}$ is definitely a rational number. Also $\dfrac{a}{a-1}$ cannot be equal to 1. Therefore $\dfrac{a}{a-1} \in Q_1$ and so it is the left inverse of a.

Also $a * b=a+b-ab=b+a-ba=b * a$.

$\therefore$ the set Q_1 of all rational numbers except 1 is an infinite abelian group with respect to the given composition.

(iii) Proceed as in part (ii). Here 0 is the identity element and the inverse of a is $-\dfrac{a}{a+1}$ which exists since $a+1 \neq 0$ and $-\dfrac{a}{a+1} \neq -1$.

Example 3. *Prove that the set of all* $m \times n$ *matrices having their elements as integers (rational or real or complex numbers) is an infinite abelian group with respect to addition of matrices.*

Solution. Let M be the set of all $m \times n$ matrices with their elements as real numbers.

Closure Property. If $A \in M$, $B \in M$, then $A+B \in M$. The reason is that $A+B$ is also a matrix of the type $m \times n$ and the elements of the matrix $A+B$ are also real numbers since the sum of two real numbers is also a real number. Therefore M is closed with respect to addition of matrices.

Associativity. We know that addition of matrices is associative.

Existence of left Identity. If O be the null matrix of the type $m \times n$ then $O \in M$. Also if $A \in M$, we have $O+A=A$.

$\therefore$ The null matrix O is the left identity.

Existence of left Inverse. If $A \in M$, then $-A \in M$ where $-A$ is the matrix whose elements are the negative of the corresponding elements of A. Also $-A+A=O=$the left identity.

$\therefore$ $-A$ is the left inverse of A.

Also the addition of matrices is commutative and the set M is an infinite set.

$\therefore$ $(M, +)$ is an infinite abelian group.

Example 4. *Show that the set of all $n \times n$ non-singular matrices having their elements as rational (real or complex) numbers is an infinite non-abelian group with respect to matrix multiplication.*

(**Berhampur 1977 ; Meerut 77**)

Solution. A matrix A is said to be non-singular if $| A |$, *i e.*, the determinant of the matrix A is not equal to zero.

Let M be the set of all $n \times n$ non-singular matrices with their elements as rational numbers.

Closure Property. We have $A \in M \Rightarrow A$ is of the type $n \times n$, the elements of A are all rational numbers and $| A | \neq 0$. Similarly let $B \in M$. Now AB will be a matrix of the type $n \times n$, the elements of AB will all be rational numbers. Also $| AB |=| A | \; | B |$. Since $| A | \neq 0$ and $| B | \neq 0$, therefore $| AB | \neq 0$. Thus $AB \in M$. Therefore M is closed with respect to multiplication of matrices.

Associativity. Multiplication of matrices is associative.

Existence of left Identity. If I be the unit matrix of the type $n \times n$, then the elements of I are all rational numbers. Also $| I |=1$ *i.e.*, $\neq 0$. Therefore $I \in M$. If $A \in M$, we have $IA=A$. Therefore I is the left identity.

Existence of left Inverse. We know that every non-singular matrix is inversible. Therefore if $A \in M$, there exists a non-singular matrix $A^{-1}=\dfrac{1}{|A|}$ (Adj. A) with elements as rational numbers such that $A^{-1}A=I=$the left identity.

We know that multiplication of matrices is not in general commutative.

$\therefore$ M is an infinite non-abelian group with respect to multiplication of matrices.

Note. If M is the set of all $n\times n$ non-singular matrices with their elements as integers, then M is not a group with respect to matrix multiplication. The reason is that all such matrices are not inversible. For example the non-singular matrix $\begin{bmatrix}1 & 2\\ 3 & 2\end{bmatrix}$ is not inversible.

We have $\begin{bmatrix}-\frac{1}{2} & \frac{1}{2}\\ \frac{3}{4} & -\frac{1}{4}\end{bmatrix}\begin{bmatrix}1 & 2\\ 3 & 2\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=$the identity element.

Since the elements of the matrix $\begin{bmatrix}-\frac{1}{2} & \frac{1}{2}\\ \frac{3}{4} & -\frac{1}{4}\end{bmatrix}$ are not all integers, therefore this matrix does not belong to the set M.

Example 5. *Show that the set of matrices*

$$A_\alpha=\begin{bmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{bmatrix}$$

where α is a real number, forms a group under matrix multiplication.

(Kanpur 1969 ; Allahabad 66)

Solution. Let G denote the set of matrices

$$A_\alpha=\begin{bmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{bmatrix},\text{ where }\alpha\in\mathbf{R}.$$

Here $\mathbf{R}$ is the set of real numbers.

To prove that G is a group with respect to matrix multiplication.

Closure Property. Let A_α, A_β be any two elements of G while $\alpha,\beta\in\mathbf{R}$. We have

$$A_\alpha A_\beta=\begin{bmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{bmatrix}\begin{bmatrix}\cos\beta & -\sin\beta\\ \sin\beta & \cos\beta\end{bmatrix}$$

$$=\begin{bmatrix}\cos(\alpha+\beta) & -\sin(\alpha+\beta)\\ \sin(\alpha+\beta) & \cos(\alpha+\beta)\end{bmatrix}=A_{\alpha+\beta}\in G,\text{ since }\alpha+\beta\in\mathbf{R}.$$

Associativity. Matrix multiplication is associative.

Existence of left Identity. Since $0\in\mathbf{R}$, therefore

$A_0=\begin{bmatrix}\cos 0 & -\sin 0\\ \sin 0 & \cos 0\end{bmatrix}\in G$. If A_α be any member of G, then

$$A_0A_\alpha=A_{0+\alpha} \qquad [\because A_\alpha A_\beta=A_{\alpha+\beta}]$$

$$=A_\alpha.$$

$\therefore$ A_0 is the left identity.

Existence of left inverse. Let $A_\alpha \in G$. Then $A_{(-\alpha)} \in G$ because $\alpha \in \mathbf{R} \Rightarrow -\alpha \in \mathbf{R}$. Now

$$A_{(-\alpha)}\, A_\alpha = A_{(-\alpha)+\alpha} \qquad [\because\ A_\alpha A_\beta = A_{\alpha+\beta}]$$
$$= A_0 = \text{the left identity.}$$

$\therefore$ $A_{(-\alpha)}$ is the left inverse of A_α.

Thus each element of G possesses left inverse. Hence G is a group under matrix multiplication.

Example 6. *Let S be any non-empty set and let A (S) be the set of all one-to-one mappings of the set S onto itself. Then show that A (S) is a group with respect to composite of mappings as the composition.* **(Banaras 1962)**

Is it an abelian group ?

Solution. Let f, g be any two elements of A (S). Then f and g are both one-to-one mappings of S onto itself. By the definition of composite of two functions f and g denoted by fg, fg is a mapping from S to S given by

$$(fg)\,(x) = f\,[g\,(x)]\ \forall\ x \in S.$$

Closure Property. We know that if f, g are two one-to-one mappings of S onto itself, then the composite mapping fg is also a one-to-one mapping of S onto itself. Consequently $fg \in A$ (S), $\forall f, g \in A$ (S). Therefore $A(S)$ is closed with respect to composite composition.

Associativity. Let f, g, h be any three elements of $A(S)$. Then $(fg)\,h$ and $f\,(gh)$ are both mappings from S to S. For every $x \in S$, we have

$$[(fg)\,h]\,(x) = (fg)\,[h\,(x)] = f\,[g\,(h(x))] = f\,[(gh)\,(x)] = [f\,(gh)]\,(x).$$

$\therefore$ $(fg)\,h = f\,(gh)$, by def. of equality of two mappings.

Existence of left Identity. Let e be the identity function on S, *i.e.*, $e(x) = x\ \forall\ x \in S$. Then the mapping e is obviously one-one onto. Therefore $e \in A$ (S). If f is any element of A (S), then $\forall\ x \in S$, we have

$$(ef)\,(x) = e\,[f\,(x)]$$
$$= f(x), \text{ since } e \text{ is identity mapping on } S.$$

Therefore $ef = f$. Thus $e \in A$ (S) is the left identity.

Existence of left Inverse. Let $f \in A$ (S). Then f is a one-to-one mapping of S onto itself. Let y be any element of S. Since f is onto S, therefore $y \in S \Rightarrow \exists\ x \in S$ such that $f(x) = y$. Also x determined in this way is a unique element of S because f is one-one.

Now let us define a mapping $f^{-1} : S \to S$ such that

$$f^{-1}(y)=x \text{ if and only if } f(x)=y.$$

Then f^{-1} is also a one-to-one mapping of S onto itself. Therefore $f^{-1} \in A(S)$. Let us compute $f^{-1}f$ which is a mapping from S to S. Given $x \in S$, let $y=f(x)$. Then by definition of f^{-1}, we have $f^{-1}(y)=x$.

$$\therefore \quad (f^{-1}f)(x)=f^{-1}[f(x)]=f^{-1}(y)=x.$$

Thus $f^{-1}f$ is the identity mapping of S *i.e.*, $f^{-1}f=e$. Therefore f^{-1} is the left inverse of f.

Hence $A(S)$ is a group with respect to composite composition.

If the set S has only one element, then the set $A(S)$ has only one element, and every group of order 1 is abelian. If the set S has two elements, then the set $A(S)$ has also two elements and every group of order 2 is abelian. But if the set S has more than two elements, then we shall show that the group $A(S)$ is non-abelian. Let x_1, x_2, x_3 be three distinct elements in S. Define the mapping $f : S \to S$ by $f(x_1)=x_2$, $f(x_2)=x_3$, $f(x_3)=x_1$, $f(y)=y$ for any $y \in S$ different from x_1, x_2, x_3. Also define the mapping $g : S \to S$ by

$$g(x_2)=x_3,\ g(x_3)=x_2 \text{ and } g\ y=y$$

for any $y \in S$ different from x_2, x_3. Obviously both f and g are one-to-one mappings of S onto itself. Thus both f and g are in $A(S)$. We have

$$(gf)(x_1)=g[f(x_1)]=g(x_2)=x_3.$$

Also $(fg)(x_1)=f[g(x_1)]=f(x_1)=x_2$.

Thus $(gf)(x_1) \neq (fg)(x_1)$. Therefore $gf \neq fg$.

Hence in this case $A(S)$ is a non-abelian group.

Example 7. *Determine as to whether the following sets together with the indicated composition constitute groups :*

(*i*) *The set S of all ordered pairs (a, b) of real numbers for which $a \neq 0$ with respect to the operation $\times$ defined by*

$$(a, b) \times (c, d)=(ac, bc+d).$$

(I.A.S. 1969 ; Madras 74 ; Kerala 70)

(*ii*) *The set P (X) of all subsets of a non-empty set X, under the composition $*$ defined by the relation*

$$A * B=A \cup B \ \forall\ A \subseteq X, B \subseteq X. \quad \textbf{(Rajasthan 1966)}$$

Solution. (i).

Closure Property. Let (a, b) and (c, d) be any two members of S. Then $a \neq 0$ and $c \neq 0$. Therefore $ac \neq 0$. Consequently $(a, b) \times (c, d)=(ac, bc+d)$ is also a member of S. Hence S is closed with respect to the given composition.

Associativity. Let (a, b), (c, d) and (e, f) be any three members of S. Then

$$\begin{aligned}[(a, b)\times(c, d)]\times(e, f) &= (ac, bc+d)\times(e, f)\\ &= ([ac]\, e, [bc+d]\, e+f)\\ &= (ace, bce+de+f).\end{aligned}$$

Also $$\begin{aligned}(a, b)\times[(c, d)\times(e, f)] &= (a, b)\times(ce, de+f)\\ &= (a\,[ce], b\,[ce]+de+f)\\ &= (ace, bce+de+f).\end{aligned}$$

Hence the given composition $\times$ is associative.

Existence of Left Identity. Suppose (x, y) is an element of S such that $(x, y)\times(a, b)=(a, b)$ $\forall$ $(a, b) \in S$.

Then $(xa, ya+b)=(a, b)$.

Hence $xa=a$ and $ya+b=b$.

These give $x=1$ and $y=0$. [Note that $a\neq 0$]

Therefore $(1, 0)$ is the left identity.

Existence of Left inverse. Let (a, b) be any member of S. Let (x, y) be a member of S such that $(x, y)\times(a, b)=(1, 0)$.

Then $(xa, ya+b)=(1, 0)$. Hence $xa=1$, $ya+b=0$.

These give $x=1/a$, $y=-b/a$.

Since $a\neq 0$, therefore x and y are real numbers.

Also $x=\frac{1}{a}\neq 0$. Thus $\left(\frac{1}{a}, -\frac{b}{a}\right)$ is the left inverse of (a, b).

Hence S is a group.

Note. In the above group, we have

$$(a, b)\times(c, d)=(ac, bc+d)$$

and $$(c, d)\times(a, b)=(ca, da+b).$$

Thus, in general, $(a, b)\times(c, d)\neq(c, d)\times(a, b)$ *i.e.*, the composition is not commutative and hence the group is not abelian.

(ii) **Closure property.** Let A and B be any two members of $P(X)$. Then $A \subseteq X$, $B \subseteq X$. We have $A * B=A \cup B$ which is also a subset of X. Thus $A * B$ is also a member of $P(X)$. Therefore $P(X)$ is closed with respect to the given operation.

Associativity. We know that the union of sets is an associative operation.

Existence of Left identity. The empty set $\varnothing$ is a subset of X. Therefore $\varnothing$ is a member of $P(X)$. If A is any member of $P(X)$, we have $\varnothing * A=\varnothing \cup A=A$. Therefore $\varnothing$ is the left identity.

Existence of left inverse. Let A be any non-empty member of $P(X)$ *i.e.*, let $A \subseteq X$ and $A \neq \emptyset$. Now for every member S of $P(X)$, we have $S * A = S \cup A \neq \emptyset$. Therefore no member S of $P(X)$ can be the left inverse of A. Hence $P(X)$ is not a group under the composition of union.

Example 8. *Let G be a set of elements on which an algebraic operation* $\times$ *is defined such that* $a \times b \in G$ *for all* $a, b \in G$. *Prove that G is an abelian group for this operation if the following postulates are satisfied* :

(*i*) $(a \times b) \times c = a \times (c \times b)$ *for all* $a, b, c \in G$;

(*ii*) *There exists a left identity* $e \in G$ *such that* $e \times a = a$ *for all* $a \in G$; *and*

(*iii*) *Corresponding to every element* $a \in G$, *there exists a left inverse* $a^{-1} \in G$ *such that* $a^{-1} \times a = e$. **(Kerala 1970)**

Solution. Let a, b be any two elements of G. By given postulate (ii) left identity e is an element of G. Applying given postulate (i) for the elements e, a and b of G, we have

$$(e \times a) \times b = e \times (b \times a)$$

$\Rightarrow a \times b = b \times a$ [$\because$ e is the left identity means $e \times a = a$ and $e \times (b \times a) = b \times a$]

Thus we have $a \times b = b \times a$ for all $a, b \in G$. Hence the operation $\times$ is commutative.

Now let $a, b, c \in G$. Then by (i), we have

$(a \times b) \times c = a \times (c \times b)$

$= a \times (b \times c)$. [$\because$ $b \times c = c \times b$ as just proved]

Hence the operation $\times$ is associative. Thus G is an abelian group for the operation $\times$.

Example 9. *Prove that in a matrix group under multiplication, either all the matrices are non-singular or all are singular.*

Show that the set of all matrices of the form $\begin{bmatrix} x & x \\ x & x \end{bmatrix}$, *where x is a non-zero real number, is a group of singular matrices for multiplication. Find the identity and inverse of an element.*

(Madras 1962)

Solution. Let M be a matrix group under multiplication. Let E be the identity element. Then

$$AE = A \;\forall\; A \in M. \qquad \ldots(1)$$

If E is a singular matrix, then the equation (1) is not satisfied by an element $A \in M$ if A is a non-singular matrix. The reason is that if A is non-singular and E is singular, then AE is singular and so it cannot be equal to a non-singular matrix A. Therefore if E

is a singular matrix, then every matrix $A \in M$ must also be singular.

Now suppose E is a non-singular matrix. Let $A \in M$ and let A be singular. Then there exists no matrix $B \in M$ such that $AB = E$. [Note that E is non-singular while AB is singular]. Thus A does not possess inverse. This contradicts the hypothesis that M is a group. Therefore if E is a non-singular matrix, then every matrix $A \in M$ must also be non-singular. Hence M cannot be a group if it contains both singular as well as non-singular matrices.

Second Part. Let $M = \left\{ \begin{bmatrix} x & x \\ x & x \end{bmatrix} : x \text{ is a non-zero real number} \right\}$.

We have $\begin{vmatrix} x & x \\ x & x \end{vmatrix} = 0$. So $\begin{bmatrix} x & x \\ x & x \end{bmatrix}$ is a singular matrix.

Now we shall show that M is a group under multiplication.

Closure Property. Let $A = \begin{bmatrix} x & x \\ x & x \end{bmatrix}$, $B = \begin{bmatrix} y & y \\ y & y \end{bmatrix}$ be any two elements of M. Then x, y are non-zero real numbers. We have

$$AB = \begin{bmatrix} 2xy & 2xy \\ 2xy & 2xy \end{bmatrix}.$$

Since $2xy$ is also a non-zero real number, therefore $AB \in M$.

Associativity. Matrix multiplication is associative.

Existence of left identity. Let $E = \begin{bmatrix} e & e \\ e & e \end{bmatrix} \in M$ be such that $EA = A \;\forall\; A \in M$. Let $A = \begin{bmatrix} x & x \\ x & x \end{bmatrix} \in M$. Then

$$\begin{aligned} EA = A &\Rightarrow \begin{bmatrix} e & e \\ e & e \end{bmatrix} \begin{bmatrix} x & x \\ x & x \end{bmatrix} = \begin{bmatrix} x & x \\ x & x \end{bmatrix} \\ &\Rightarrow \begin{bmatrix} 2ex & 2ex \\ 2ex & 2ex \end{bmatrix} = \begin{bmatrix} x & x \\ x & x \end{bmatrix} \\ &\Rightarrow 2ex = x \\ &\Rightarrow e = \tfrac{1}{2}, \text{ since } x \neq 0. \end{aligned}$$

Thus $E = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \in M$ and is such that $EA = A \;\forall\; A \in M$.

Therefore E is the left identity.

Existence of left inverse. Let $A = \begin{bmatrix} x & x \\ x & x \end{bmatrix}$ be an arbitrary element of M. Suppose $B = \begin{bmatrix} y & y \\ y & y \end{bmatrix} \in M$ is such that $BA = E$.

Then we have

$$\begin{bmatrix} 2xy & 2xy \\ 2xy & 2xy \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

This gives $2xy=\frac{1}{2}$ or $y=1/4x$. Thus $B=\begin{bmatrix}1/4x & 1/4x\\ 1/4x & 1/4x\end{bmatrix}\in M$ is such that $BA=E$. Therefore B is the left inverse of A.

Hence M is a group.

Identity$=\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2}\end{bmatrix}$, $\begin{bmatrix}x & x\\ x & x\end{bmatrix}^{-1}=\begin{bmatrix}1/4x & 1/4x\\ 1/4x & 1/4x\end{bmatrix}$.

Example 10. *If* $G=\left\{\begin{bmatrix}a & 0\\ 0 & 0\end{bmatrix} : a \text{ is any non-zero real number}\right\}$, *show that G is a commutative group under matrix multiplication.*

(Madras 1983)

Solution. Let $A=\begin{bmatrix}a & 0\\ 0 & 0\end{bmatrix}$, $B=\begin{bmatrix}b & 0\\ 0 & 0\end{bmatrix}$ be any two elements of G. Then a, b are non-zero real numbers. We have

$$AB=\begin{bmatrix}ab & 0\\ 0 & 0\end{bmatrix}. \text{ Also } BA=\begin{bmatrix}ba & 0\\ 0 & 0\end{bmatrix}.$$

Since ab is also a non-zero real number, therefore $AB \in G$. Thus G is closed for matrix multiplication. Also $AB=BA$ because $ab=ba$. Therefore multiplication on G is commutative. Also we know that matrix multiplication is associative.

Existence of left identity. Let $E=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$. Then $E \in G$ and we have

$$EA=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}a & 0\\ 0 & 0\end{bmatrix}=\begin{bmatrix}a & 0\\ 0 & 0\end{bmatrix}=A, \quad \forall\ A \in G.$$

Therefore E is the left identity.

Existence of left inverse. Let $A=\begin{bmatrix}a & 0\\ 0 & 0\end{bmatrix}$ be an arbitrary element of G. Then a is a non-zero real number and so $1/a$ is also a non-zero real number. Therefore $B=\begin{bmatrix}1/a & 0\\ 0 & 0\end{bmatrix}$ is also an element of G. We have

$$BA=\begin{bmatrix}1/a & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}a & 0\\ 0 & 0\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}=E.$$

$\therefore$ B is the left inverse of A.

Hence G is a commutative group for matrix multiplication.

§ 7. Composition table for finite sets.

A binary composition in a finite set can be shown in a tabular form known as **Composition table.**

Suppose $S=\{a_1, a_2, a_3, \ldots, a_n\}$ is a finite set having n elements. Suppose there is a composition in S denoted multiplicatively. We write the elements of the set S in a horizontal row as well as in a vertical column. The element $a_i a_j$ associated to the ordered pair (a_i, a_j) is placed at the intersection of the row headed by a_i and the column headed by a_j.

An Important Remark. The composition table for a finite group contains each element exactly once in each of its rows and columns. For example, let $a_i a_j$ and $a_i a_k$ be any two elements of the ith row. Then $a_i a_j = a_i a_k \Rightarrow a_j = a_k$ (by left cancellation law). But this is a contradiction since a_j, a_k are different. Hence the result.

Some Examples On Finite Groups

Example 1. *Show that the four fourth roots of unity namely*

$$1, -1, i, -i$$

form a group with respect to multiplication. (Meerut 1976)

Solution. Let $G=\{1, -1, i, -i\}$. To show that multiplication is a composition in G, we form the composition table.

multiplication	1	-1	i	$-i$
1	1	-1	i	$-i$
-1	-1	1	$-i$	i
i	i	$-i$	-1	1
$-i$	$-i$	i	1	-1

Now we make the following conclusions :

1. Closure property. Since all the entries in the composition table are elements of the set G, G is closed with respect to multiplication. Therefore multiplication is a binary operation on G.

2. Associativity. The elements of G are all complex numbers and the multiplication of complex numbers is associative.

3. Existence of left identity. From the composition table we see that the row headed by the element 1 just coincides with the top row of the composition table. Thus we have

$$1\,(1)=1,\ 1\,(-1)=-1,\ 1\,(i)=i,\ 1\,(-i)=-i.$$

In other words we have $1 \in G$ and $1a=a \; \forall \; a \in G$.

$\therefore$ 1 is the left identity.

4. Existence of left inverse. We know that the identity element is its own inverse. Therefore the left inverse of 1 is 1. From the composition table we see that in the column headed by -1, the left identity 1 occurs in the row headed by -1 *i.e.*, $(-1)(-1)=1$. Therefore -1 is the left inverse of -1. Also we see that in the column headed by i the left identity 1 occurs in the row headed by $-i$. Thus $(-i)\,i=1$. Therefore $-i$ is the left inverse of i. Finally in the column headed by $-i$ the left identity 1 occurs in the row headed by i. Thus $i\,(-i)=1$. Therefore i is the left inverse of $-i$. Thus each element of G possesses left inverse.

Hence G is a group with respect to multiplication. The number of elements in the set G is 4. Also the multiplication of complex numbers is commutative. Therefore G is an abelian group of order 4 with respect to multiplication.

Note. To see that the composition is commutative we observe from the composition table that the entries in the first, second, third and fourth rows of the table just coincide with the corresponding entries in the first, second, third and fourth columns. Therefore we have $ab=ba \;\forall\; a, b \in G$ and the composition is commutative.

Example 2. *Show that the set $G=\{1, \omega, \omega^2\}$, where ω is an imaginary cube root of unity is a group with respect to multiplication.*
(Agra 1971; Madurai 78; Meerut 84; Kumayun 77; Rohilkhand 80)

Solution. We form the composition table :

Multiplication	1	ω	ω^2
1	1	ω	ω^2
ω	ω	ω^2	1
ω^2	ω^2	1	ω

Note that
$\omega\omega^2=\omega^3=1$ and
$\omega^2\omega^2=\omega^4=\omega^3\omega$
$=1\omega$
$=\omega$

1. Since all the entries in the composition table are elements of the set G, therefore G is closed with respect to multiplication.

2. The elements of G are all complex numbers and we know that multiplication of complex numbers is associative.

3. From the composition table we see that

$$1\,(1)=1,\; 1\,(\omega)=\omega=\omega\,(1),\; 1\,(\omega^2)=\omega^2=\omega^2\,(1).$$

Therefore 1 is the identity element.

4. The inverses of 1, ω, ω^2 are 1, ω^2, ω respectively.

5. The multiplication of complex numbers is commutative. The number of elements in the set G is 3.

Hence G is a finite abelian group of order 3.

Example 3. *Show that the set of six transformations* $f_1, f_2, f_3, f_4, f_5, f_6$ *on the set of complex numbers defined by*

$$f_1(z)=z,\ f_2(z)=\frac{1}{z},\ f_3(z)=1-z,\ f_4(z)=\frac{z}{z-1},$$

$$f_5(z)=\frac{1}{1-z},\ f_6(z)=\frac{z-1}{z}$$

forms a finite non-abelian group of order six with respeet to the composition known as composite of two functions or product of two functions. **(Sambalpur 1977)**

Solution. Let $G=\{f_1, f_2, f_3, f_4, f_5, f_6\}$.

Suppose we denote multiplicatively the composition known as the **composite** or **product** of two functions. If $f: A\to B$ and $g: B\to C$ then by definition $(gf): A\to C$ such that $(gf)(x)=g[f(x)]\ \forall x\in A$. The function gf is called the composite of the functions g and f. We prepare the composition table as follows:

Since the function f_1 is the identity function, therefore

$$f_1 f_1=f_1,\ f_1 f_2=f_2=f_2 f_1,\ f_1 f_3=f_3=f_3 f_1,$$
$$f_1 f_4=f_4=f_4 f_1,\ f_1 f_5=f_5=f_5 f_1,\ f_1 f_6=f_6=f_6 f_1.$$

Now $(f_2 f_2)(z)=f_2[f_2(z)]=f_2\left(\frac{1}{z}\right)=\frac{1}{(1/z)}=z=f_1(z)$.

$\therefore\ f_2 f_2=f_1$ $\quad$ [$\because$ if $f: A\to B$ and $g: A\to B$, then $f=g$ iff $f(x)=g(x)\ \forall x\in A$]

$(f_2 f_3)(z)=f_2[f_3(z)]=f_2(1-z)=\frac{1}{1-z}=f_5(z)$. Therefore $f_2 f_3=f_5$.

$(f_2 f_4)(z)=f_2[f_4(z)]=f_2\left(\frac{z}{z-1}\right)=\frac{z-1}{z}=f_6(z)$. Therefore $f_2 f_4=f_6$.

$(f_2 f_5)(z)=f_2[f_5(z)]=f_2\left(\frac{1}{1-z}\right)=1-z=f_3(z)$. Therefore $f_2 f_5=f_3$.

Similarly calculating the other products we get the composition table as given below:

Composition of two functions	f_1	f_2	f_3	f_4	f_5	f_6
f_1	f_1	f_2	f_3	f_4	f_5	f_6
f_2	f_2	f_1	f_5	f_6	f_3	f_4
f_3	f_3	f_6	f_1	f_5	f_4	f_2
f_4	f_4	f_5	f_6	f_1	f_2	f_3
f_5	f_5	f_4	f_2	f_3	f_6	f_1
f_6	f_6	f_3	f_4	f_2	f_1	

We make the following observations :

1. All the entries in the composition table are elements of the set G, therefore G is closed with respect to the given composition.

We know that the composite of functions is an associative composition *i.e.*, if $f: A \to B$, $g: B \to C$, $h: C \to D$,
then $h(gf)=(hg)f$.

3. The identity function f_1 is the identity element.

4. Each function possesses inverse. Thus
$f_1^{-1}=f_1, f_2^{-1}=f_2, f_3^{-1}=f_3, f_4^{-1}=f_4, f_5^{-1}=f_6, f_6^{-1}=f_5.$

5. The composition is not commutative since
$f_2 f_3=f_5$ and $f_3 f_2=f_6$. Thus $f_2 f_3 \neq f_3 f_2$.

The set G contains 6 elements.

Hence G is a finite non-abelian group of order six with respect to the composite composition.

Note. Here we see in the composition table that the entries in the second row do not coincide with the corresponding entries in the second column. Thus $f_2 f_3 \neq f_3 f_2$. Therefore the composition is not commutative.

Example 4. *Prove that the set of all n nth roots of unity forms a finite abelian group of order n with respect to multiplication.*

(Banaras 1969; Meerut 81; Allahabad 70)

Solution. We have $(1)^{1/n}=(1+0i)^{1/n}=(\cos 0+i \sin 0)^{1/n}$
$=(\cos 2r\pi+i \sin 2r\pi)^{1/n}$, where r is any integer
$=\cos \dfrac{2r\pi}{n}+i \sin \dfrac{2r\pi}{n}$ [by De Moivre's theorem]
$=e^{(i2r\pi)/n}$

Putting $r=0, 1, 2, \ldots\ldots, n-1$ we get the n nth roots of unity.

$\therefore$ If G is the set of the n nth roots of unity, then

$$G=\{1, e^{2\pi i/n}, e^{4\pi i/n}, e^{6\pi i/n}, \ldots, e^{[2(n-1)\pi i]/n}\},$$

or $$G=\{1, \omega, \omega^2, \omega^3, \ldots, \omega^{n-1}\}, \text{ where } \omega=e^{2\pi i/n}.$$

Closure property. Let $a, b \in G$. Then both a and b are nth roots of unity. Therefore $a^n=1$ and $b^n=1$.

Now $(ab)^n=a^n b^n=1\times 1=1$.

$\therefore$ ab is also an nth root of unity and so $ab \in G$.

Thus $a, b \in G \Rightarrow ab \in G$. Therefore G is closed with respect to multiplication.

Associativity. The elements of G are all complex numbers and the multiplication of complex numbers is associative.

Existence of left identity. We have $1 \in G$ and $1\,a=a \;\forall\, a \in G$. Therefore 1 is the left identity.

Existence of left inverse. The left inverse of 1 is 1.

If ω^r, $1 \leqslant r \leqslant n-1$, is any other element of G, then ω^{n-r} is also an element of G. We have

$$\omega^{n-r}\,\omega^r=\omega^n=1 \qquad [\because \ \omega \text{ is an } n\text{th root of unity}]$$

$\therefore$ ω^{n-r} is the left inverse of ω^r.

Further the multiplication of complex numbers is **commutative.**

$\therefore$ the set of n nth roots of unity is a finite abelian group of order n with respect to the operation of multiplication.

Example 5. Quaternion Group. Let $T=\{\pm 1, \pm i, \pm j, \pm k\}$. Define a multiplicative binary operation on T by setting $i^2=j^2=k^2=-1$ and $ij=-ji=k, jk=-kj=i$, and $ki=-ik=j$. It can be easily shown that for this binary operation T is a non-abelian group. T is called a *Quaternion Group* and its order is 8. It can be easily verified that the set G consisting of the following eight matrices

$$\pm\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \pm\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \pm\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}, \pm\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

forms a quaternion group under the operation of matrix multiplication.

Exercises

1. Show that the set $G=\{1, -1\}$ is a finite abelian group of order 2 under multiplication as composition.

2. (i) Distinguish between an abelian and a non-abelian group. Give an example of each. **(I. A. S. 1970; Kumayon 78)**

(ii) In group theory, prove that the left axioms imply the right axioms. **(Sambalpur 1977)**

3. Show that the set

$$G=\{\ldots, 3^{-4}, 3^{-3}, 3^{-2}, 3^{-1}, 1, 3, 3^2, 3^3, 3^4, \ldots\}$$

forms an infinite abelian group with respect to multiplication.

4. Show that the set of all positive rational numbers (real numbers) forms an abelian group with respect to multiplication of numbers. **(Meerut 1972)**

5. Show that the set I of all integers is an abelian group with operation * defined by $a*b=a+b+2$.

6. Prove that the set of rational numbers of the form $\frac{m}{2^n}$ (m, n integers) is a group under addition.

7. Show that the set G of all square matrices $[a_{ij}]_{n\times n}$ such that det $[a_{ij}]=\pm 1$ is a group under matrix multiplication. Show also that those matrices in G for which det $[a_{ij}]=1$ form a group.

8. Show that the set of all matrices $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, a and b being non-zero reals, is a group under matrix multiplication.

9. Show that the set of all matrices $\begin{bmatrix} c^a & b \\ 0 & c^{-a} \end{bmatrix}$, a and b real, is a group under matrix multiplication, c being a positive constant.

10. Show that the set of all matrices $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, where a and b are real numbers not both equal to zero, is a group under matrix multiplication. **(Madras 1962)**

11. Show that the set of matrices

$$A_\alpha=\begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

where α is a real number, forms a group under matrix multiplication. **(I.A.S. 1975)**

12. Show that the set of four transformations f_1, f_2, f_3, f_4 on the set of complex numbers defined by

$$f_1(z)=z,\ f_2(z)=-z,\ f_3(z)=1/z,\ f_4(z)=-1/z,$$

forms a finite abelian group with respect to the composite composition. **(Banaras 1963)**

13. Show that the set of complex numbers z with $|z|=1$ is not a group under the operation * denoted by

$$z_1 * z_2=|z_1|.z_2.$$

14. Show that the four matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

form a multiplicative group. **(Rajasthan 1975)**

15. If z denotes any complex number, then show that the set of all bilinear transformations

$$f(z)=\frac{az+b}{cz+d},\ ad-bc\neq 0,\ a,\ b,\ c,\ d\in \mathbf{C}$$

is an infinite non-abelian group.

16. Prove that the set of all transformations $f(z)=\dfrac{az+b}{cz+d}$ with $ad-bc=1$ is a group (a, b, c and d are complex numbers).

17. If G is a group and $a\in G$ is such that $aa=a$, prove that $a=e$.

18. If every element of a group is its own inverse, show that the group must be abelian.

19. Show that the composition table for a finite group contains each group element once and only once in each of its rows and columns.

20. Forming the composition table for the multiplicative group $\{e, a, b\}$ of order 3, show that every group of order 3 must be abelian.

21. Let G be a non-empty set closed under an associative product, which in addition satisfies :

(i) **Existence of right identity** *i.e.*, there exists an element $e\in G$ such that $ae=a\ \forall\ a\in G$.

(ii) **Existence of right inverse** *i.e.*, $a\in G\Rightarrow$ there exists an element $a^{-1}\in G$ such that $aa^{-1}=e$.

Prove that G must be a group under this product.

(Nagarjuna 1979)

[**Hint.** First prove that with the given postulates the right cancellation law must hold. Then show that the right identity is also the left identity and the right inverse of an element is also its left inverse.]

22. Let M be the set of all 2×2 matrices of the form $\begin{bmatrix} x & y \\ x & y \end{bmatrix}$, where $x, y\in \mathbf{R}$ with $x+y\neq 0$. Show that M is a semi-group under matrix multiplication. Further show that M has a left identity and each element of M has a right inverse. Is M a group ? (**Ans.** No.)

23. Let $\mathbf{R}_0$ be the set of all real numbers except zero. Define a binary operation $*$ on $\mathbf{R}_0$ by $a * b=|a|\,b$ where $|a|$ denotes the absolute value of a. Show that

(i) $*$ is associative on $\mathbf{R}_0$.

(ii) There exists a left identity for $*$, and a right inverse for each element in $\mathbf{R}_0$.

Is $(\mathbf{R}_0, *)$ a group ? (**Ans.** No.)

[**Hint.** Note that 1 and -1 are both left identities.]

§ 8. Addition Modulo m. We shall now define a new type of addition known as "addition modulo m" and written as $a+_m b$ where a and b are any integers and m is a fixed positive integer.

By definition, we have

$$a+_m b=r,\ 0 \leqslant r < m$$

where r is the least non-negative remainder when $a+b$ (*i.e.*, the ordinary sum of a and b) is divided by m.

For example $15+_5 7=2$, since $15+7=22=4(5)+2$ *i.e.*, 2 is the least non-negative remainder when $15+7$ is divided by 5. Similarly $5+_6 3=2$; $5+_7 2=0$; $9+_{12} 3=0$; $4+_3 2=0$; $3+_3 1=1$. Also $-23+_7 3=1$, since $-23+3=-20=(-3)\,7+1$.

Thus to find $a+_m b$, we add a and b in the ordinary way and then from the sum, we remove integral multiples of m in such a way that the remainder r is either 0 or a positive integer less than m.

When a and b are two integers such that $a-b$ is divisible by a fixed positive integer m, then we write

$$a \equiv b \pmod{m}$$

which is read as "a is congruent to b modulo m".

Thus $a \equiv b \pmod{m}$ iff $a-b$ is divisible by m. For example $9 \equiv 4 \pmod{5}$ since $9-4$ is divisible by 5. $13 \equiv 3 \pmod{5}$; $5 \equiv 5 \pmod{5}$; $9 \equiv 3 \pmod{6}$; $16 \equiv 4 \pmod{6}$; $-20 \equiv 4 \pmod{6}$.

It can be easily seen that if $a \equiv b \pmod{m}$, then $a+_m c=b+_m c$.

For $a \equiv b \pmod{m} \Rightarrow a-b$ is divisible by m

$\Rightarrow a-b=km$ for some integer k

$\Rightarrow a=b+km$.

Now $a+_m c=(b+km)+_m c$

$=$least non-negative remainder when $(b+km)+c$ is divided by m

$=$least non-negative remainder when $b+c$ is divided by m

$=b+_m c$.

For example $5 \equiv 1 \pmod{4}$ and we have $5+_4 2=1+_4 2=3$.

Also it is obvious that $a+_m b \equiv a+b \pmod{m}$.

For example $9+_4 5=2$ and $9+5=14$. Now $14 \equiv 2 \pmod{4}$.

Finally $\qquad a+_m b=b+_m a$.

§ 9. Multiplication modulo p. We shall now define a new type of multiplication known as "multiplication modulo p" and written as $a \times_p b$ where a and b are any integers and p is a fixed positive integer. By definition, we have

$$a \times_p b = r,\ 0 \leqslant r < p,$$

where r is the least non-negative remainder when ab (*i.e.*, the ordinary product of a and b) is divided by p. For example $8 \times_5 3 = 4$ since $8 \times 3 = 24 = 4\ (5) + 4$.

Also $4 \times_7 2 = 1$ since $4 \times 2 = 8 = 1\ (7) + 1$.

It can be easily seen that if $a \equiv b$ *(mod* p*), then* $a \times_p c = b \times_p c$.

Also it is obvious that $a \times_p b \equiv ab$ (mod p). For example $5 \times_3 4 = 2$ and $5 \times 4 = 20$. Now we see that $2 \equiv 20$ (mod 3) since $2 - 20$ is divisible by 3.

Theorem 1. Additive group of integers modulo m. *The set* $G = \{0, 1, 2, \ldots, m-1\}$ *of first* m *non-negative integers is a group, the composition being addition reduced modulo* m.

Proof. We have by definition of addition modulo m,

$$a +_m b = r$$

where r is the least non-negative remainder when the ordinary sum $a+b$ is divided by m. Obviously $0 \leqslant r \leqslant (m-1)$. Therefore for all $a, b \in G$, we have $a +_m b \in G$ and thus G is closed with respect to the composition addition modulo m.

Associativity. Let a, b, c be any arbitrary elements of G. Then

$a +_m (b +_m c) = a +_m (b + c)$ $\quad [\because\ b +_m c \equiv b + c\ (\text{mod}\ m)]$

$=$ least non-negative remainder when $a + (b + c)$ is divided by m

$=$ least non-negative remainder when $(a + b) + c$ is divided by m,

since $a + (b + c) = (a + b) + c$

$= (a + b) +_m c$ $\quad$ [by def. of $+_m$]

$= (a +_m b) +_m c$ $\quad [\because\ a + b \equiv a +_m b\ (\text{mod}\ m)]$

$\therefore$ '$+_m$' is an associative composition.

Existence of Identity Element. We have $0 \in G$. Also if a is any element of G, then $0 +_m a = a = a +_m 0$. Therefore 0 is the identity element.

Existence of inverse. The inverse of 0 is 0 itself. If $r \in G$ and $r \neq 0$, then $m - r \in G$. Also $(m - r) +_m r = 0 = r +_m (m - r)$. Therefore $m - r$ is the inverse of r.

Also, the composition '$+_m$' is commutative, since

$a +_m b =$ least non-negative remainder when $a + b$ is divided by m

$=$ least non-negative remainder when $b + a$ is divided by m

$= b +_m a$.

The set G contains m elements.

$\therefore$ $(G, +_m)$ is a finite abelian group of order m.

Note. If we exclude zero from the above set, it will not form a group.

Theorem 2. Multiplicative group of integers modulo p where p is prime. *The set G of $(p-1)$ integers $1, 2, 3, \ldots, p-1$, p being prime, is a finite abelian group of order $p-1$, the composition being multiplication modulo p.*

Proof. Let $G=\{1, 2, 3, \ldots, p-1\}$ where p is prime.

[A non-zero integer p is called a prime integer if it is neither 1 nor -1 and if its only divisors are $1, -1, p, -p$. The first 10 positive primes are 2, 3, 5, 7, 11, 13, 17, 19, 23, 29].

Let a and b be any elements of G. Then $1 \leqslant a \leqslant p-1$, $1 \leqslant b \leqslant p-1$. Now by definition, $a \times_p b=r$ where r is the least non-negative remainder when the ordinary product ab is divided by p. Since p is prime, therefore ab is not exactly divisible by p. Therefore r cannot be zero and we shall have $1 \leqslant r \leqslant p-1$. Thus $a \times_p b \in G \ \forall \ a, b \in G$. Hence the closure axiom is satisfied.

Associativity. Let a, b, c be any arbitrary elements of G.

Then $a \times_p (b \times_p c)=a \times_p (bc)$ $\quad [\because \ b \times_p c \equiv bc \pmod{p}]$

$=$least non-negative remainder when $a\,(bc)$ is divided by p

$=$least non-negative remainder when $(ab)\,c$ is divided by p

$=(ab) \times_p c$

$=(a \times_p b) \times_p c$ $\quad [\because \ ab \equiv a \times_p b \pmod{p}]$

$\therefore \ \times_p$ is an associative composition.

Existence of left identity. We have $1 \in G$. Also if a is any element of G, then $1 \times_p a=a$. Therefore 1 is the left identity.

Existence of left inverse. Let s be any member of G. Then $1 \leqslant s \leqslant p-1$. Consider the following $p-1$ products :

$$1 \times_p s,\ 2 \times_p s,\ 3 \times_p s,\ \ldots,\ (p-1) \times_p s.$$

All these are elements of G. Also no two of these can be equal as shown below :

Let i and j be two unequal integers such that

$$1 \leqslant i \leqslant p-1,\ 1 \leqslant j \leqslant p-1 \text{ and } i>j.$$

Then $\qquad i \times_p s=j \times_p s$

$\Rightarrow$ is and js leave the same least non-negative remainder when divided by p

$\Rightarrow$ $is-js$ is divisible by $p \Rightarrow (i-j)\,s$ is divisible by p.

Since $1 \leqslant (i-j) < p-1$; $1 \leqslant s \leqslant p-1$ and p is prime therefore $(i-j)\,s$ cannot be divisible by p.

$$\therefore \ i \times_p s \neq j \times_p s.$$

Thus $1 \times_p s, 2 \times_p s, \ldots, (p-1) \times_p s$ are $p-1$ distinct elements of the set G. Therefore one of these elements must be equal to 1.

Let $s' \times_p s = 1$. Then s' is the left inverse of s.

Finally the composition '$\times_p$' is commutative, since

$a \times_p b$ = least non-negative remainder when ab is divided by p

= least non-negative remainder when ba is divided by p

$= b \times_p a$.

$\therefore$ $(G, \times_p)$ is a finite abelian group of order $p-1$.

Note 1. Suppose in the set G, p is not prime but p is composite. Then $\exists$ two integers a and b such that $1 < a \leqslant p-1$, $1 < b \leqslant p-1$ and $ab = p$. Therefore $a \times_p b = 0$ and $0 \notin G$. Therefore G will not be closed with respect to the composition multiplication modulo p. Then $(G, \times_p)$ will not be a group.

Note 2. If we include 0 in the set G, then also for this composition G will not be a group. The reason is that the inverse of 0 will not exist.

Example 1. *Prove that the set* $G = \{0, 1, 2, 3, 4, 5\}$ *is a finite abelian group of order* 6 *with respect to addition modulo* 6.

Solution. Let us form the composition table.

$+_6$	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

We see that all the entries in the composition table are elements of the set G. Therefore G is closed with respect to addition modulo 6 *i.e.*, $+_6$.

The composition '$+_6$' **is associative.** If a, b, c are any three elements of G, then

$a +_6 (b +_6 c) = a +_6 (b + c)$ $\quad [\because \; b +_6 c \equiv b + c \pmod 6]$

= least non-negative remainder when $a + (b + c)$ is divided by 6

= least non-negative remainder when $(a + b) + c$ is divided by 6

$= (a + b) +_6 c = (a +_6 b) +_6 c$. $\quad [\because \; a + b \equiv a +_6 b \pmod 6]$

Existence of Identity. We have $0 \in G$. If a is any element of G, then from the composition table we see that

$$0 +_6 a = a = a +_6 0.$$

Therefore 0 is the identity element.

Existence of Inverse. From the table we see that the inverses of 0, 1, 2, 3, 4, 5 are 0, 5, 4, 3, 2, 1 respectively. For example $4 +_6 2 = 0 = 2 +_6 4$ implies 4 is the inverse of 2.

The composition is commutative as the corresponding rows and columns in the composition table are identical. The number of elements in the set G is 6.

$\therefore$ $(G, +_6)$ is a finite abelian group of order 6.

Example 2. *Prove that the set* $G = \{1, 2, 3, 4, 5, 6\}$ *is a finite abelian group of order* 6 *with respect to multiplication modulo* 7.

Solution. Let us form the composition table :

$\times_7$	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	1	3	5
3	3	6	2	5	1	4
4	4	1	5	2	6	3
5	5	3	1	6	4	2
6	6	5	4	3	2	1

We see that all the entries in the composition table are elements of the set G. Therefore G is closed with respect to multiplication modulo 7 *i.e.*, $\times_7$.

The composition '$\times_7$' is associative. If a, b, c are any three elements of G, then

$a \times_7 (b \times_7 c) = a \times_7 (bc)$ $\quad$ $[\because\ b \times_7 c \equiv bc \pmod 7]$
= least non-negative remainder when $a(bc)$ is divided by 7
= least non-negative remainder when $(ab)c$ is divided by 7
$= (ab) \times_7 c = (a \times_7 b) \times_7 c$ $\quad$ $[\because\ a \times_7 b \equiv ab \pmod 7]$

Existence of Identity. We have $1 \in G$. If a is any element of G, then from the composition table we see that $1 \times_7 a = a = a \times_7 1$.

$\therefore$ 1 is the identity element.

Existence of Inverse. From the table we see that the inverses of 1, 2, 3, 4, 5, 6, are 1, 4, 5, 2, 3, 6 respectively. For example $3 \times_7 5 = 1 = 5 \times_7 3$ implies 3 is the inverse of 5.

The composition is commutative as the corresponding rows and columns in the composition table are identical. The set G has 6 elements. Hence $(G, \times_7)$ is a finite abelian group of order 6.

§ 10. Residue classes of the set of integers.

Some properties of integers. Let $\mathbf{I} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ be the set of integers.

Division algorithm. *Let* $a, b \in \mathbf{I}$ *with* $b \neq 0$. *Then we can divide a by b to get a non-negative remainder r which is smaller in size than b. In other words if* $a, 0 \neq b \in \mathbf{I}$, *then there exist integers q and r such that* $a = qb + r$ *where* $0 \leqslant r < |b|$.

For example let $-15, -9$ be two integers. Then we can write $-15 = 2(-9) + 3$. Here $0 \leqslant 3 < |-9|$.

Divisibility in the set of integers.

Definition. *Let* $a, 0 \neq b \in \mathbf{I}$. *We say that a is divisible by b if* $a = bm$ *where m is some integer.*

If a is divisible by b, then we also say that b is a divisor of a. In symbols we shall write it as $b \mid a$. It will be read as 'b is a divisor of a'. If $a \mid b$, $b \mid c$, then $a \mid c$. Also if $a \mid b$, $a \mid d$, then $a \mid (b+d)$. Further if $a \mid b$, then $a \mid (bk)$ where k is any integer.

Associates. Definition. *Two non-zero integers a and b are said to be associates if* $a \mid b$ *and* $b \mid a$. If a and b are associates, then we must have $a = \pm b$.

Greatest common divisor. Definition. *Let a and b be any two integers. Then the positive integer c is said to be the greatest common divisor of a and b if*

(*i*) $c \mid a$ *and* $c \mid b$.

(*ii*) *whenever* $d \mid a$ *and* $d \mid b$, *then* $d \mid c$.

The greatest common divisor of the integers a and b will be symbolically denoted by (a, b).

Existence and uniqueness of greatest common divisor. *If a and b are integers, not both* 0, *then they have a unique greatest common divisor, say, c. Also we can find integers x and y such that*

$$c = xa + yb.$$

Relatively prime integers. Definition. *Two integers a and b are said to be relatively prime if* $(a, b) = 1$ *i.e., if their greatest common divisor is* 1. For example 3 and 5 are relatively prime integers, 1 and 8 are relatively prime integers. If a and b are relatively prime integers then $(a, b)=1$. Therefore we can find integers x and y such that $1=xa+yb$.

Prime Integers. Definition. *An integer p is said to be a prime integer if* $p \neq 0$, $p \neq \pm 1$, *and the only divisors of p are* ± 1, $\pm p$. For example ± 2, ± 3, ± 5, ± 7, ± 11,... are prime integers.

Three very important results.

(*i*) *Let a, b, c be three integers such that a and b are relatively prime and* $a \mid bc$. *Then we must have* $a \mid c$.

(*ii*) *Let p be a prime integer and a, b be any two integers such that* $p \mid ab$. *Then we must have* $p \mid a$ *or* $p \mid b$.

(*iii*) *Let a be any integer greater than* 1. *Then a can be uniquely expressed (except for order) as the product of a finite number of positive primes.* For example $18=2 \times 3 \times 3=3 \times 2 \times 3$.

Relation of '*congruence modulo m*' in the set of integers.

Definition. *Let m be any fixed positive integer. Then an integer a is said to be congruent to another integer b modulo m if* $m \mid (a-b)$ *i.e., if* $(a-b)$ *is divisible by m.* Symbolically we write $a \equiv b \pmod{m}$. It will be read as "*a* is congruent to *b* modulo *m*".

For example $13 \equiv 3 \pmod{5}$ because $13-3=10$ which is divisible by 5. Further $17 \equiv -3 \pmod{5}$ because $17-(-3)=20$ which is divisible by 5.

Theorem 1. *Show that "congruence modulo m" is an equivalence relation in the set of integers. Further show that this equivalence relation has m distinct equivalence classes.* **(Allahabad 1979)**

Proof. Let **I** be the set of integers. If m is any positive integer then we say that $a \equiv b \pmod{m}$ if $m \mid (a-b)$. We shall prove that this defines an equivalence relation in the set **I**.

Reflexivity. Let a be any integer. Then $a-a=0$ and $m \mid 0$. Thus $a \equiv a \pmod{m}$ $\forall$ $a \in \mathbf{I}$. Therefore the relation is reflexive.

Symmetry. Let $a, b \in \mathbf{I}$ be such that $a \equiv b \pmod{m}$. Then we have $m \mid (a-b) \Rightarrow m \mid -(a-b) \Rightarrow m \mid (b-a) \Rightarrow b \equiv a \pmod{m}$. Thus $a \equiv b \pmod{m} \Rightarrow b \equiv a \pmod{m}$. Therefore the relation is symmetric.

Transitivity. Let $a, b, c \in \mathbf{I}$ be such that $a \equiv b \pmod{m}$, $b \equiv c \pmod{m}$. Then we have $m \mid (a-b)$ and $m \mid (b-c)$ $\Rightarrow m \mid \{(a-b)+(b-c)\} \Rightarrow m \mid (a-c) \Rightarrow a \equiv c \pmod{m}$. Thus $a \equiv b \pmod{m}$, $b \equiv c \pmod{m} \Rightarrow a \equiv c \pmod{m}$. Therefore the relation is transitive.

Hence this is an equivalence relation.

Consequently it will partition **I** into disjoint equivalence classes called *residue classes modulo m or congruence classes modulo m.*

If $a \in \mathrm{I}$, then the residue class $\bar{a}$ or $\{a\}$ or $[a]$

$=\{x : x \in \mathrm{I}$ and $x-a$ is divisible by $m\}$.

If $b \in \mathrm{I}$, then the residue class $[b]$

$=\{y : y \in \mathrm{I}$ and $y-b$ is divisible by $m\}$.

Obviously $[a]=[b]$ if and only if $a \equiv b \pmod{m}$ *i.e.*, if and only if $m \mid (a-b)$. Thus $[a]=[a+m]=[a+2m]$ and so on. Similarly $[1]=[1+m]=[1+2m]$ and so on. Also $[0]=[m]=[2m]=[-m]$ and so on.

Now let I_m be the set of all residue classes of integers modulo m. This set is also called the set of integers mod m. [Some authors denote this set by J_m]. We shall prove that the set I_m has m distinct elements. We claim that

$$\mathrm{I}_m=\{[0], [1], [2], \ldots, [m-1]\}.$$

First we shall show that the m residue classes $[0], [1], \ldots, [m-1]$ are all distinct. Let $0 \leqslant i < m$, $0 \leqslant j < m$ and $j > i$.

Then $[i]=[j] \Rightarrow i \equiv j \pmod{m}$

$\Rightarrow i-j$ is divisible by $m \Rightarrow j-i$ is divisible by m.

But according to our assumption $j-i$ is a positive integer less than m. So it cannot be divisible by m. Therefore $[i] \neq [j]$ and thus $[0], [1], \ldots, [m-1]$ are all distinct.

Now we shall show that if a is any integer, then the residue class $[a]$ is equal to one of the residue classes $[0], [1], \ldots, [m-1]$. By division algorithm, we have

$a=km+r$, where $k, r \in \mathrm{I}$ and $0 \leqslant r < m$

$\Rightarrow a-r=km \Rightarrow a-r$ is divisible by $m \Rightarrow a \equiv r \pmod{m} \Rightarrow [a]=[r]$.

Since $0 \leqslant r \leqslant m-1$, therefore the residue class $[a]=[r]$ is one of the residue classes $[0], [1], \ldots, [m-1]$.

Hence the set I_m has m distinct elements.

Note. The residue class $[0]$ is called the *zero residue class*. We have $[a]=[0]$ if and only if $m \mid a$. The set of integers mod m will have $m-1$ distinct non-zero residue classes.

Addition of residue classes. Definition. *If* $[a], [b] \in \mathrm{I}_m$, *then we define* $[a]+[b]=[a+b]$. Here '+' on the L.H.S. stands for addition of residue classes and '+' on the R.H.S. stands for addition of integers.

Since $a, b \in \mathrm{I} \Rightarrow a+b \in \mathrm{I}$, therefore $[a+b]$ is also a residue class *i.e.*, $[a+b] \in \mathrm{I}_m$. Now we know that $[a]=[a+m]=[a+2m]$, and so on. Thus a residue class can be represented in several ways. Therefore we must show that our addition of residue classes is well defined *i.e.*, it is independent of the representation of any residue class. For this we are to show that if $[a]=[c]$ and $[b]=[d]$, then $[a]+[b]=[c]+[d]$.

We have $[a]=[c] \Rightarrow m \mid (a-c)$. Also $[b]=[d] \Rightarrow m \mid (b-d)$.

Now $m \mid (a-c)$ and $m \mid (b-d)$

$\Rightarrow m \mid \{(a-c)+(b-d)\} \Rightarrow m \mid \{(a+b)-(c+d)\}$

$\Rightarrow a+b \equiv c+d \pmod m \Rightarrow [a+b]=[c+d] \Rightarrow [a]+[b]=[c]+[d]$.

Thus $[a]=[c]$ and $[b]=[d] \Rightarrow [a]+[b]=[c]+[d]$.

Hence our addition of residue classes is well-defined.

Multiplication of residue classes. Definition. *If* $[a], [b] \in \mathbf{I}_m$, *then we define* $[a]\,[b]=[ab]$.

Since $a, b \in \mathbf{I} \Rightarrow ab \in \mathbf{I}$, therefore $[ab]$ is also a residue class *i.e.*, $[ab] \in \mathbf{I}_m$. But we must show that our multiplication of residue classes is **well-defined.** For this we are to show that if $[a]=[c]$ and $[b]=[d]$, then $[a]\,[b]=[c]\,[d]$.

We have $[a]=[c] \Rightarrow a \equiv c \pmod m$

$\Rightarrow a-c$ is divisible by $m \Rightarrow b\,(a-c)$ is divisible by m.

Also $[b]=[d] \Rightarrow b-d$ is divisible by $m \Rightarrow c\,(b-d)$ is divisible by m.

$\therefore$ $[a]=[c]$ and $[b]=[d] \Rightarrow \{b\,(a-c)+c\,(b-d)\}$ is divisible by m

$\Rightarrow ab-cd$ is divisible by m

$\Rightarrow ab \equiv cd \pmod m \Rightarrow [ab]=[cd]$

$\Rightarrow [a]\,[b]=[c]\,[d]$.

Hence our multiplication of residue classes is well defined.

Theorem 2. *The set of residue classes modulo m is an abelian group of order m with respect to addition of residue classes.*

(Rajasthan 1974 ; Meerut 79)

Proof. Let $\mathbf{I}=\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ be the set of integers. If $a \in \mathbf{I}$, then $[a]$ is a residue class modulo m of $\mathbf{I}$ if

$[a]=\{x : x \in \mathbf{I}$, and $x-a$ is divisible by $m\}$.

Let $\mathbf{I}_m$ be the set of all residue classes of $\mathbf{I}$ mod m *i.e.*, $\mathbf{I}_m=\{[a] : a \in \mathbf{I}\}$. We have $[a]=[b] \Leftrightarrow a \equiv b \pmod m$. The set $\mathbf{I}_m$ has m distinct elements $[0], [1], [2], \ldots, [m-1]$. Thus we have

$$\mathbf{I}_m=\{[0], [1], [2], \ldots, [m-1]\}.$$

If a and b are any two integers, then we define the addition of residue classes $[a]$ and $[b]$ as follows :

$$[a]+[b]=[a+b].$$

If $[a]=[c]$ and $[b]=[d]$, then it can be easily seen that

$$[a]+[b]=[c]+[d].$$

Therefore our addition of residue classes is well defined.

Now we shall show that I_m is a group with respect to addition of residue classes.

Closure property. If $[a], [b] \in I_m$, then by definition $[a]+[b] = [a+b]$. Since $a+b$ is an integer, therefore $[a+b] \in I_m$. Thus I_m is closed with respect to addition of residue classes.

Associativity. Let $[a], [b], [c]$ be any three elements of I_m. Then $[a]+([b]+[c])$

$= [a]+[b+c]$ [by def. of addition of residue classes]

$= [a+(b+c)]$ [by def. of addition of residue classes]

$= [(a+b)+c]$ [$\because$ addition of integers is associative]

$= [a+b]+[c] = ([a]+[b])+[c]$.

Commutativity. Let $[a], [b] \in I_m$. Then

$$[a]+[b] = [a+b] = [b+a] = [b]+[a].$$

Existence of identity. We have $[0] \in I_m$. If $[a] \in I_m$, then we have $[0]+[a] = [0+a] = [a] = [a]+[0]$. Therefore the residue class $[0]$ is the identity element.

Existence of inverse. Let $[a] \in I_m$ be arbitrary. Since $a \in I \Rightarrow -a \in I$, therefore $[-a]$ is also an element of I_m. We have $[-a]+[a] = [-a+a] = [0] = [a]+[-a]$. Thus $[-a]$ is the inverse of $[a]$.

Thus I_m is an abelian group with respect to addition of residue classes. Since the number of distinct elements in I_m is m, therefore the order of this group is m.

Note. If $[r] \in I_m$ and $0 \leqslant r < m$, then the inverse of $[r]$ is $[m-r]$. We have $[r]+[m-r] = [r+(m-r)] = [m] = [0]$. Note that

$$m \mid m \Rightarrow [m] = [0].$$

Theorem 3. *The set of non-zero residue classes modulo a prime integer p forms an abelian group of order $p-1$ with respect to multiplication of residue classes.* (Meerut 1981; Jiwaji 78; Gorakhpur 70)

Proof. Let $I = \{..., -3, -2, -1, 0, 1, 2, 3, ...\}$ be the set of integers. If $a \in I$, then $[a]$ is a residue class modulo p of I if

$$[a] = \{x : x \in I \text{ and } x-a \text{ is divisible by } p\}.$$

We have $[a]=[b]$ if and only if $a-b$ is divisible by p.

The residue class $[0]$ is called the zero residue class. We have $[a]=[0]$ if and only if a is divisible by p.

Let I_p be the set of residue classes of I mod p. Then I_p has p distinct elements *i.e.*, $[0], [1], [2], ..., [p-1]$. The residue class $[0]$ is the zero residue class. Therefore if G is the set of all non-zero residue classes modulo p, then G has $p-1$ distinct elements and we have $\quad G=\{[1], [2], \ldots, [p-1]\}$

or $\quad G=\{[a] : a \in \mathbf{I}$ and a is not divisible by $p\}$.

If a and b are any two integers, then we define the multiplication of residue classes $[a]$ and $[b]$ as follows :

$$[a]\ [b]=[ab].$$

If $[a]=[c]$ and $[b]=[d]$, then it can be easily seen that $[a]\ [b]=[c]\ [d]$. Therefore our multiplication of residue classes is well defined.

Now we shall show that G is a group with respect to multiplication of residue classes.

Closure property. Let $[a]$, $[b]$ be any two elements of G. Then a and b are integers each not divisible by p. By definition, we have $\quad [a]\ [b]=[ab]$.

Since $ab \in \mathbf{I}$, therefore $[ab]$ is also a residue class mod p. But we must show that $[ab]$ is a non-zero residue class *i.e.*, p is not a divisor of ab.

We have $p \mid ab \Rightarrow p \mid a$ or $p \mid b$, since p is a prime integer.

But neither a nor b is divisible by p. Hence ab is not divisible by p. Thus $[ab] \neq [0]$. Consequently $[ab] \in G$ and G is closed with respect to multiplication of residue classes.

Associativity. If $[a]$, $[b]$, $[c]$ be any three residue classes, then

$$\begin{aligned} [a]\ ([b]\ [c]) &= [a]\ [bc] && \text{(by def.)} \\ &= [a\ (bc)] && \text{(by def.)} \\ &= [(ab)\ c] && (\because \text{ multiplication of integers is associative}) \\ &= [ab]\ [c] = ([a]\ [b])\ [c]. \end{aligned}$$

$\therefore$ multiplication of residue classes is associative.

Commutativity. If $[a]$, $[b]$ be any two residue classes, then

$$[a]\ [b]=[ab]=[ba]=[b]\ [a].$$

$\therefore$ multiplication of residue classes is commutative.

Existence of identity. Since the integer 1 is not divisible by the prime integer p, therefore $[1] \neq [0]$. Thus $[1] \in G$. If $[a]$ be any element of G, we have

$$[1]\ [a]=[1a]=[a]=[a]\ [1].$$

Thus the residue class $[1]$ is the identity element.

Existence of inverse. Let $[a] \in G$ *i.e.*, let $[a]$ be any non-zero residue class. Then a is not divisible by p. Consider the $p-1$ products

$$[1]\ [a], [2]\ [a], \ldots, [p-1]\ [a].$$

By closure property all these products are elements of G. We claim that all these are distinct elements of G.

Let i and j be two unequal integers, such that

$$1 \leqslant i \leqslant p-1,\ 1 \leqslant j \leqslant p-1 \text{ and } i > j.$$

Then $[i][a]=[j][a] \Rightarrow [ia]=[ja]$

$\Rightarrow ia-ja$ is divisible by $p \Rightarrow p \mid (i-j)\, a$

$\Rightarrow p \mid a$ or $p \mid (i-j)$, since p is a prime integer

$\Rightarrow p \mid (i-j)$, since p is not a divisor of a.

But according to our assumption $i-j$ is a positive integer less than p. So it cannot be divisible by p. Therefore $[i]\,[a] \neq [j]\,[a]$. Thus $[1]\,[a], [2]\,[a], \ldots, [p-1]\,[a]$ are the $p-1$ distinct elements of G placed in some order. Therefore one of these elements must be equal to $[1]$. Let $[b]\,[a]=[1]=[a]\,[b]$. Then $[b]$ is the inverse of $[a]$.

Thus G is an abelian group with respect to multiplication of residue classes. Since the number of distinct elements in G is $p-1$, therefore the order of this group is $p-1$.

Theorem 4. *The set of non-zero residue classes modulo a composite positive integer m is not a group with respect to multiplication of residue classes.* **(Banaras 1966)**

Proof. Let G be the set of non-zero residue classes modulo a composite positive integer m. Since m is a composite integer, therefore let $m=ab$ where $1 < a < m$, $1 < b < m$. Now $1 < a < m \Rightarrow m$ is not a divisor of $a \Rightarrow [a] \neq [0] \Rightarrow [a] \in G$. Similarly $[b] \neq [0]$ and therefore $[b] \in G$.

Now $ab=m \Rightarrow [ab]=[m]$

$\Rightarrow [a]\,[b]=[0]$, since $[m]=[0]$

$\Rightarrow [a]\,[b] \notin G$.

Thus $[a], [b] \in G \Rightarrow [a]\,[b] \notin G$. Therefore G is not closed with respect to multiplication of residue classes. Hence G is not a group with respect to multiplication of residue classes.

§ 11. An alternative set of postulates for a group. Theorem. *A set G with a binary composition denoted multiplicatively is a group if*

(*i*) *the composition is associative.*

(*ii*) *for every pair of elements* $a, b \in G$, *the equations*

$$ax=b \text{ and } ya=b$$

have solutions in G. **(Delhi Hons. 1970 ; Rajasthan 76 ; Poona 73 ; Jiwaji 78; Andhra 77; Marathwada 72; Osmania 72)**

Proof. In order to prove that a set G equipped with a composition satisfying conditions (i) and (ii) is a group, we should show that the left identity exists and each element of G possesses left inverse.

It is given that for every pair of elements $a, b \in G$ the equation $ya=b$ has a solution in G. Therefore if $a \in G$, then taking $b=a$, we see that there exists an element, $e \in G$ such that

$$ea=a. \qquad \ldots(1)$$

Suppose now b is any arbitrary element of G. Since $a \in G$, therefore from (ii) there exists $x \in G$ such that

$$ax=b. \qquad \ldots(2)$$

Now
$$\begin{aligned} eb &= e\,(ax) && [\because\ b=ax] \\ &= (ea)\,x && [\text{by associativity}] \\ &= ax && [\text{from } (1)] \\ &= b. && [\text{from } (2)] \end{aligned}$$

Thus $\exists\ e \in G$ such that $eb=b\ \forall\ b \in G$. Therefore e is the left identity.

Suppose now a is any arbitrary element of G. Since $e \in G$, therefore taking $b=e$ in the given condition (ii), we see that the equation $ya=e$ has a solution in G Let $c \in G$ be such that $ca=e$. Then c is the left inverse of a. Therefore each element of G possesses left inverse.

Since left identity exists and each element possesses left inverse, therefore the left identity will also be the right identity and the left inverse of any element will also be its right inverse.

Hence G is a group for the given composition if the postulates (i) and (ii) are satisfied.

Example 1. *Prove that a finite set G with a composition denoted multiplicatively is a group if the composition is associative and the right and left cancellation laws hold in G i.e.,*

$$ax=bx \Rightarrow a=b,\ xa=xb \Rightarrow a=b.$$

Show further that the validity of the cancellation laws does not characterise infinite groups but characterises finite groups only.

(Nagarjuna 1980 ; Allahabad 79 ; Meerut 70)

Solution. Suppose the set G has n distinct elements

$$a_1, a_2, a_3, \ldots, a_n.$$

Let a be any one of these elements. Then the n products

$$a_1a, a_2a, a_3a, \ldots, a_na$$

are all elements of G. Also these n elements are all distinct. For $a_ia=a_ja$ where $a_i, a_j \in G \Rightarrow a_i=a_j$ (by right cancellation law).

$\therefore$ if $a_i \neq a_j$, then $a_i\, a \neq a_j\, a$.

Thus $a_1a, a_2a, a_3a, \ldots\ldots, a_na$ are nothing but the n elements of G placed in some order.

Therefore if b is any element of G, then one of these n elements will be equal to b.

Therefore if a, b are any two elements of G, there exists an element, say $c \in G$ such that $ca=b$. In other words, the equation

$$ya=b$$

has a solution in G for every pair of elements $a, b \in G$.

Similarly by forming the products

$$aa_1, aa_2, aa_3, \ldots\ldots, aa_n$$

and by using the left cancellation law, we can show that the equation $$ax=b$$
has a solution in G for every pair of elements $a, b \in G$.

Hence the postulates (i) and (ii) of § 11 are satisfied. Therefore G is a finite group of order n.

However, an infinite set will not necessarily form a group even if the composition in the set is associative and both the cancellation laws hold. For example, the set N of all natural numbers

$$1, 2, 3, 4, \ldots$$

is not a group for multiplication even though

(i) N is closed with respect to multiplication.

(ii) Multiplication in N is associative.

(iii) Both the cancellation laws hold good for multiplication in N.

Euler's ϕ function. Definition. *The Euler ϕ-function, $\phi(n)$, is defined for all positive integers n by :*

$\phi(1)=1$; *for $n > 1$, $\phi(n)=$ the number of positive integers less than n and relatively prime to n.*

Thus $\phi(6)=2$ because the positive integers less than 6 and relatively prime to 6 are 5, 1 and their number is 2. Similarly $\phi(8)=4$ because the positive integers less than 8 and relatively prime to 8 are 7, 5, 3, 1 and their number is 4. If p is a positive prime integer, then obviously $\phi(p)=p-1$.

Example 2. *Let n be any positive integer. For any integer a, let $[a]$ denote the residue class of the set of integers mod n. Let $G=\{[a] : a$ is an integer relatively prime to $n\}$. Prove that with respect to multiplication of residue classes G is a group of order $\phi(n)$, where ϕ is the Euler ϕ-function.*

Solution. Let $I=\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ be the set of integers and let n be any positive integer. For any integer a, let

$[a]=\{x : x \in I \text{ and } x \equiv a \pmod{n}\}$.

Let $G=\{[a] : a$ is an integer relatively prime to n *i.e.*, $(a, n)=1\}$.

First we shall show that the number of distinct elements in G is equal to $\phi(n)$.

Case 1. $n=1$. In this case if a is any integer, then

$a-1$ is divisible by $1 \Rightarrow a \equiv 1 \pmod{1} \Rightarrow [a]=[1]$.

Thus in this case each residue class $[a]$ is equal to the residue class $[1]$. Also 1 is relatively prime to 1.

Therefore in this case the set G has only one distinct element. Thus the number of distinct elements in $G=1=\phi(1)$.

Case 2. $n > 1$.

In this case the number of positive integers less than n and relatively prime to $n=\phi(n)=k$, say. Let $a_1, a_2, \ldots, a_k$ be the positive integers less than n and relatively prime to n. We shall show that $[a_1], [a_2], \ldots, [a_k]$ are the only distinct elements of G.

First we shall show that the residue classes $[a_1], \ldots, [a_k]$ are all distinct. Let i, j, be two positive integers less than n and relatively prime to n and let $i > j$. We have

$[i]=[j] \Rightarrow i \equiv j \pmod{n} \Rightarrow i-j$ is divisible by n.

But according to our assumption $i-j$ is a positive integer less than n. Therefore $i-j$ cannot be divisible by n. Thus $[i] \neq [j]$. Hence the residue classes $[a_1], \ldots, [a_k]$ are all distinct.

Now we shall show that if a is any integer relatively prime to n, then the residue class $[a]$ is equal to one of the residue classes $[a_1], \ldots, [a_k]$. Since a is any integer and n is a positive integer, therefore by division algorithm there exist integers q and r such that $\quad a=nq+r$ where $0 \leqslant r < n$.

We have $a-r=nq \Rightarrow a-r$ is divisible by $n \Rightarrow a \equiv r \pmod{n}$ $\Rightarrow [a]=[r]$.

Now it remains to show that r is relatively prime to n. Suppose r is not relatively prime to n. Let $(r, n)=s$ where $s>1$. Then $s \mid n$ and $s \mid r$. Therefore $s \mid (nq+r)$ *i.e.*, $s \mid a$. Now $s \mid a$ and $s \mid n$ implies that a and n are not relatively prime to each other which is a contradiction. Hence r must be relatively prime to n. Since $n>1$, therefore r cannot be 0. Hence r is a positive integer less than n and relatively prime to n. Thus $[a]$ is one of the residue classes $[a_1], \ldots, [a_k]$.

Thus the number of distinct elements in $G=k=\phi(n)$. Thus G is a finite set whether $n=1$ or $n>1$. Now we shall show that G is a group with respect to multiplication of residue classes.

Closure property. Let $[a]$, $[b]$ be any two elements of G. Then a and b are integers relatively prime to n. By definition of multiplication of residue classes, we have

$$[a]\,[b]=[ab].$$

Now $a, b \in \mathrm{I} \Rightarrow ab \in \mathrm{I}$. We shall show that ab is relatively prime to n. Suppose ab is not relatively prime to n. Let $(n, ab)=t$ where $t > 1$. The integer t can be expressed as the product of a finite number of prime integers. Let the prime integer p be a factor of t. Then $p \mid t$. Also $t \mid ab$. Therefore $p \mid ab$. Since p is prime, therefore $p \mid ab \Rightarrow p \mid a$ or $p \mid b$.

Let $p \mid a$. Now $p \mid t$ and $t \mid n$. Therefore $p \mid n$. Thus $p \mid a$ and $p \mid n$. Therefore a is not relatively prime to n. This is a contradiction Hence ab must be relatively prime to n.

$\therefore$ $[ab] \in G$ and G is closed with respect to multiplication of residue classes.

Associativity. If $[a]$, $[b]$, $[c]$ are any three residue classes, then $[a]\,([b]\,[c])=[a]\,[bc]=[a\,(bc)]=[(ab)\,c]=[ab]\,[c]=([a]\,[b])\,[c]$.

Thus multiplication of residue classes is associative.

Existence of cancellation laws. Let $[a]$, $[b]$, $[c]$ be any three elements of G. Then a, b, c are integers relatively prime to n.

We have $[a]\,[b]=[a]\,[c] \Rightarrow [ab]=[ac]$

$\Rightarrow ab-ac$ is divisible by $n \Rightarrow n \mid a\,(b-c)$

$\Rightarrow n \mid b-c$, since n is relatively prime to a

$\Rightarrow b \equiv c \pmod{n} \Rightarrow [b]=[c]$.

Thus the left cancellation law holds good in G.

Since the multiplication of residue classes is commutative, therefore the right cancellation law will also hold good in G.

Now G is a finite set closed with respect to multiplication of residue classes. Also this composition is associative and both the cancellation laws hold in G for this composition. Hence G is a group for this composition. The order of G is equal to $\phi\,(n)$, the number of distinct elements in G.

Since 1 is an integer relatively prime to n, therefore $[1] \in [G]$. If $[a]$ is any element of G, we have $[1]\,[a]=[1a]=[a]=[a]\,[1]$. Therefore $[1]$ is the identity element of this group.

Exercises

1. Show that the relation 'congruence modulo m' is an equivalence relation in the set of integers and has m distinct equivalence classes. **(Jabalpur 1970)**

2. Is the set $\{1, 2, 3, 4, 5\}$ a group under (i) addition modulo 6 (ii) multiplication modulo 6 ? **Ans.** (i) No. (ii) No.

3. Is the set $\{1, 2, 3, 4, 5, 6\}$ a group under addition modulo 7 ? **Ans.** No. **(Kolhapur 1973)**

4. Does the set of residue classes modulo 5 form a group with respect to addition ? **Ans.** Yes. **(Meerut 1973)**

5. Prove that the set $\{0, 1, 2, 3, 4\}$ is a finite abelian group of order 5 under addition modulo 5 as composition.

(Meerut 1973)

6. Prove that the set $\{1, 2, 3, 4\}$ is a finite abelian group of order 4 under multiplication modulo 5 as composition.

7. Show that the set $\{1, 3, 4, 5, 9\}$ is an abelian group under multiplication modulo 11 as composition. What is the order of this group ? **Ans.** 5.

8. Prove that $G=\{1, 5, 7, 11\}$ is a group under multiplication modulo 12.

9. Which of the following sets are groups under multiplication modulo 11 ?

(i) $\{1, 3, 5, 7, 8\}$, (ii) $\{1, 8\}$, (iii) $\{1, 10\}$. **Ans.** Only $\{1, 10\}$.

10. Show that the non-zero residue classes modulo a positive prime integer p form a commutative group with respect to multiplication of residue classes. **(Gorakhpur 1970)**

11. Define a semigroup and a group, and prove that a semigroup G is a group if and only if the equations $ax=b$ and $ya=b$ have solutions in G for arbitrary $a, b \in G$. **(Sagar 1965)**

12. Let $S=\{x \in \mathbf{I} : 1 \leqslant x \leqslant n$, and $(x, n)=1$ *i.e.*, x and n are relatively prime integers$\}$ Prove that S is a group with respect to multiplication modulo n as composition.

§ 12. Permutations.

Definition. *Suppose S is a finite set having n distinct elements. Then a one-one mapping of S onto itself is called a permutation of degree n.* **(Rajasthan 1977; Meerut 73)**

The number of elements in the finite set S is known as the degree of permutation.

Symbol for a permutation. Let $S=\{a_1, a_2, a_3, \ldots, a_n\}$ be a finite set having n distinct elements. If $f : S \rightarrow S$ and f is one-one onto, then f is a permutation of degree n. Let $f(a_1)=b_1$, $f(a_2)=b_2$, $f(a_3)=b_3, \ldots, f(a_n)=b_n$, where $\{b_1, b_2, \ldots, b_n\}=\{a_1, a_2, \ldots, a_n\}$ *i.e.*, $b_1, b_2, \ldots, b_n$ is nothing but some arrangement of the n elements of

S. We find it convenient to introduce a two line notation to write this permutation. In this notation, we write

$$f=\begin{pmatrix} a_1 & a_2 & a_3 \ldots a_n \\ b_1 & b_2 & b_3 \ldots b_n \end{pmatrix} \text{ i.e.,}$$

each element in the second row is the f-image of the element of the first row lying directly above it.

If $S=\{1, 2, 3, 4\}$ is a finite set having four elements, then

$$f=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}, \; g=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix}, \text{ etc.,}$$

are all permutations of degree 4. Here in the permutation f the elements 1, 2, 3, 4 have been replaced respectively by the elements 2, 4, 1, 3. Thus $f(1)=2$, $f(2)=4$, $f(3)=1$, $f(4)=3$. In other words each element in the first row is to be replaced by the element directly below it in the second row.

Equality of two permutations. *Two permutations f and g of degree n are said to be equal if we have $f(a)=g(a)$ $\forall$ $a \in S$.*

For example, if $f=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$ and $g=\begin{pmatrix} 2 & 4 & 3 & 1 \\ 3 & 1 & 4 & 2 \end{pmatrix}$

are two permutations of degree 4, then we have $f=g$. Here we see that both f and g replace 1 by 2, 2 by 3, 3 by 4 and 4 by 1.

If $f=\begin{pmatrix} a_1 & a_2 & a_3 \ldots a_n \\ b_1 & b_2 & b_3 \ldots b_n \end{pmatrix}$ is a permutation of degree n, we can write it in several ways. The interchange of columns will not change the permutation. Thus we can write

$$f=\begin{pmatrix} a_2 & a_1 & a_3 \ldots a_n \\ b_2 & b_1 & b_3 \ldots b_n \end{pmatrix}=\begin{pmatrix} a_n & a_1 \ldots a_2 \\ b_n & b_1 \ldots b_2 \end{pmatrix}=\begin{pmatrix} a_n & a_{n-1} \ldots a_2 & a_1 \\ b_n & b_{n-1} \ldots b_2 & b_1 \end{pmatrix} \text{ etc.}$$

Therefore if f and g are two permutations of degree n, then we can always write g in such a way that the first row of g coincides with the second row of f.

For example if $f=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}$ and $g=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$ are two permutations of degree 4, then by interchanging columns we can write $g=\begin{pmatrix} 2 & 4 & 1 & 3 \\ 3 & 1 & 4 & 2 \end{pmatrix}$.

Total number of distinct permutations of degree n. If S is a finite set having n distinct elements, then we shall have n ! distinct arrangements of the elements of S. Therefore there will be n ! distinct permutations of degree n. If P_n be the set consisting of all permutations of degree n, ten the hset P_n will have n ! distinct elements. **This set P_n is called the symmetric set of permutations of degree n.** Sometimes it is also denoted by S_n. Thus

$$P_n = \{ f : f \text{ is a permutation of degree } n \}.$$

The set P_3 of all permutations of degree 3 will have 3 ! *i.e.*, 6 elements. Obviously

$$P_3 = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\}.$$

Identity Permutation. *If I is a permutation of degree n such that I replaces each element by the element itself, I is called the identity permutation of degree n.*

Thus $I = \begin{pmatrix} 1 & 2 & 3 \ldots n \\ 1 & 2 & 3 \ldots n \end{pmatrix}$ or $\begin{pmatrix} a_1 & a_2 & a_3 \cdot a_n \\ a_1 & a_2 & a_3 \ldots a_n \end{pmatrix}$ or $\begin{pmatrix} b_1 & b_2 & b_3 \ldots b_n \\ b_1 & b_2 & b_3 \ldots b_n \end{pmatrix}$ is the identity permutation of degree n.

Product or Composite of two permutations. *The product or composite of two permutations f and g of degree n denoted by fg, is obtained by first carrying out the operation defined by f and then by g.*

Suppose P_n is the set of all permutations of degree n. Let

$$f = \begin{pmatrix} a_1 & a_2 & a_3 \ldots a_n \\ b_1 & b_2 & b_3 \ldots b_n \end{pmatrix} \text{ and } g = \begin{pmatrix} b_1 & b_2 & b_3 \ldots b_n \\ c_1 & c_2 & c_3 \ldots c_n \end{pmatrix}$$

be any two elements of P_n.

Here the permutation g has been written in such a way that the first row of g coincides with the second row of f. If the product of the permutations f and g is denoted multiplicatively *i.e.*, by fg, then by definition

$$fg = \begin{pmatrix} a_1 & a_2 & a_3 \ldots a_n \\ c_1 & c_2 & c_3 \ldots c_n \end{pmatrix}.$$

For, f replaces a_1 by b_1 and then g replaces b_1 by c_1 so that fg replaces a_1 by c_1. Similarly fg replaces a_2 by c_2, a_3 by $c_3, \ldots, a_n$ by c_n.

Obviously fg is also a permutation of degree n. Thus the product of two permutations of degree n is also a permutation of degree n. Therefore $fg \in P_n \ \forall f, g \in P_n$.

Example 1. *Let* $f = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$ *and* $g = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ *be two permutations of degree* 3. *Then*

$$fg = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

$$=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\begin{pmatrix}1&3&2\\2&1&3\end{pmatrix}=\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$$

and $gf=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}$

$$=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}\begin{pmatrix}2&3&1\\3&2&1\end{pmatrix}=\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}.$$

Obviously $fg\neq gf$. The reason is that fg replaces 1 by 2 while gf replaces 1 by 3. So fg cannot be equal to gf.

Thus we see that multiplication of permutations is not in general commutative.

Example 2. *Let* $f=\begin{pmatrix}1&2&3&4&5\\2&3&4&5&1\end{pmatrix}$ *and* $g=\begin{pmatrix}1&2&3&4&5\\1&2&4&5&3\end{pmatrix}$ *be two permutations of degree* 5

To obtain fg there is no necessity that we should write g in such a way that the first row of g coincides with the second row of f.

We see that f replaces 1 by 2 and g replaces 2 by 2. Therefore fg replaces 1 by 2. Again f replaces 2 by 3 and g replaces 3 by 4 therefore fg replaces 2 by 4. Further f replaces 3 by 4 and g replaces 4 by 5, therefore fg replaces 3 by 5. Proceeding in this way we get

$$fg=\begin{pmatrix}1&2&3&4&5\\2&4&5&3&1\end{pmatrix}.$$

§ 13. Groups of Permutations.

Theorem. *The set P_n of all permutations on n symbols is a finite group of order n ! with respect to composite of mappings as the operation. For $n \leqslant 2$, this group is abelian and for $n > 2$ it is always non-abelian.*

(Rajasthan 1976; Meerut 72, 73, 79, 80; Allahabad 69)

Proof. Let $S=\{a_1, a_2, \ldots, a_n\}$ be a finite set having n distinct elements. Let $f=\begin{pmatrix}a_1 & a_2 \ldots a_n\\ b_1 & b_2 \ldots b_n\end{pmatrix}$ be a permutation of degree n.

Here elements $b_1, b_2, \ldots, b_n$ of the second row are simply an arrangement of the n elements $a_1, a_2, \ldots, a_n$ of the set S.

The elements of the set S can be arranged in n ! different ways. Therefore we shall have n ! distinct permutations of degree n. If P_n be the set of all permutations of degree n, then P_n has n ! distinct elements.

Let $f=\begin{pmatrix} a_1 & a_2 \dots a_n \\ b_1 & b_2 \dots b_n \end{pmatrix}$ and $g=\begin{pmatrix} b_1 & b_2 \dots b_n \\ c_1 & c_2 \dots c_n \end{pmatrix}$ be any two permutations of degree n. Then by definition of product or composite of two permutations, denoted multiplicatively, we have

$$fg=\begin{pmatrix} a_1 & a_2 \dots a_n \\ c_1 & c_2 \dots c_n \end{pmatrix}.$$

Obviously fg is also a permutation of degree n, since $c_1, c_2, \dots, c_n$ is nothing but an arrangement of the same n elements $a_1, a_2, \dots, a_n$ of the set S. Thus $fg \in P_n \, \forall f, g \in P_n$. Therefore P_n is closed with respect to the composition known as product of two permutations.

Associativity. Permutation multiplication is associative. Let

$$f=\begin{pmatrix} a_1 & a_2 \dots a_n \\ b_1 & b_2 \dots b_n \end{pmatrix}, \; g=\begin{pmatrix} b_1 & b_2 \dots b_n \\ c_1 & c_2 \dots c_n \end{pmatrix}, \; h=\begin{pmatrix} c_1 & c_2 \dots c_n \\ d_1 & d_2 \dots d_n \end{pmatrix}$$

be any three permutations of degree n where $b_1, b_2, \dots, b_n$; $c_1, c_2, \dots, c_n$; $d_1, d_2, \dots, d_n$ are simply different arrangements of the same n elements $a_1, a_2, \dots, a_n$.

Then $(fg)=\begin{pmatrix} a_1 & a_2 \dots a_n \\ c_1 & c_2 \dots c_n \end{pmatrix}$.

$\therefore \quad (fg)\,h=\begin{pmatrix} a_1 & a_2 \dots a_n \\ c_1 & c_2 \dots c_n \end{pmatrix}\begin{pmatrix} c_1 & c_2 \dots c_n \\ d_1 & d_2 \dots d_n \end{pmatrix}=\begin{pmatrix} a_1 & a_2 \dots a_n \\ d_1 & d_2 \dots d_n \end{pmatrix}$.

Also $(gh)=\begin{pmatrix} b_1 & b_2 \dots b_n \\ d_1 & d_2 \dots d_n \end{pmatrix}$.

$\therefore \quad f\,(gh)=\begin{pmatrix} a_1 & a_2 \dots a_n \\ b_1 & b_2 \dots b_n \end{pmatrix}\begin{pmatrix} b_1 & b_2 \dots b_n \\ d_1 & d_2 \dots d_n \end{pmatrix}=\begin{pmatrix} a_1 & a_2 \dots a_n \\ d_1 & d_2 \dots d_n \end{pmatrix}$.

Thus $(fg)\,h=f\,(gh)$.

Existence of Identity. Let

$$I=\begin{pmatrix} a_1 & a_2 \dots a_n \\ a_1 & a_2 \dots a_n \end{pmatrix} \text{ or } \begin{pmatrix} b_1 & b_2 \dots b_n \\ b_1 & b_2 \dots b_n \end{pmatrix}$$

be the identity permutation of degree n. Then $I \in P_n$.

If $f=\begin{pmatrix} a_1 & a_2 \dots a_n \\ b_1 & b_2 \dots b_n \end{pmatrix}$ is any element of P_n, we have

$$fI=\begin{pmatrix} a_1 & a_2 \dots a_n \\ b_1 & b_2 \dots b_n \end{pmatrix}\begin{pmatrix} b_1 & b_2 \dots b_n \\ b_1 & b_2 \dots b_n \end{pmatrix}=\begin{pmatrix} a_1 & a_2 \dots a_n \\ b_1 & b_2 \dots b_n \end{pmatrix}=f.$$

Also $If=\begin{pmatrix} a_1 & a_2 \dots a_n \\ a_1 & a_2 \dots a_n \end{pmatrix}\begin{pmatrix} a_1 & a_2 \dots a_n \\ b_1 & b_2 \dots b_n \end{pmatrix}=\begin{pmatrix} a_1 & a_2 \dots a_n \\ b_1 & b_2 \dots b_n \end{pmatrix}=f.$

$\therefore$ Identity permutation I is the identity element.

Existence of Inverse. Let $f=\begin{pmatrix} a_1 & a_2 \dots a_n \\ b_1 & b_2 \dots b_n \end{pmatrix}$ be any element of P_n. Then $f^{-1}=\begin{pmatrix} b_1 & b_2 \dots b_n \\ a_1 & a_2 \dots a_n \end{pmatrix}$ is also an element of P_n since f^{-1} is also a permutation of degree n. We have

$$f^{-1}f=\begin{pmatrix} b_1 & b_2 \dots b_n \\ a_1 & a_2 \dots a_n \end{pmatrix}\begin{pmatrix} a_1 & a_2 \dots a_n \\ b_1 & b_2 \dots b_n \end{pmatrix}=\begin{pmatrix} b_1 & b_2 \dots b_n \\ b_1 & b_2 \dots b_n \end{pmatrix}=I.$$

Also $$ff^{-1}=\begin{pmatrix} a_1 & a_2 \dots a_n \\ b_1 & b_2 \dots b_n \end{pmatrix}\begin{pmatrix} b_1 & b_2 \dots b_n \\ a_1 & a_2 \dots a_n \end{pmatrix}=\begin{pmatrix} a_1 & a_2 \dots a_n \\ a_1 & a_2 \dots a_n \end{pmatrix}=I.$$

$\therefore$ f^{-1} is the inverse of f.

Hence P_n is a group of order n ! with respect to product of permutations as composition.

If $n=1$, the set P_n has only one element and every group of order 1 is abelian. If $n=2$, the set P_n has 2 ! *i.e.*, 2 elements and every group of order 2 is again abelian. Now we shall show that if $n > 2$, P_n is non-abelian. Let

$$f=\begin{pmatrix} 1 & 2 & 3 \dots n-1 & n \\ 2 & 3 & 4 \dots n & 1 \end{pmatrix} \text{ and } g=\begin{pmatrix} 1 & 2 & 3 & 4 \dots n-1 & n \\ 2 & 1 & 3 & 4 \dots n-1 & n \end{pmatrix}$$

be the two permutations of degree n, when $n > 2$. Then

$$fg=\begin{pmatrix} 1 & 2 & 3 \dots n-1 & n \\ 1 & 3 & 4 \dots n & 2 \end{pmatrix} \text{ and } gf=\begin{pmatrix} 1 & 2 & 3 & 4 \dots n-1 & n \\ 3 & 2 & 4 & 5 \dots n & 1 \end{pmatrix}.$$

Obviously $fg \neq gf$. Therefore P_n is non-abelian if $n > 2$.

Note 1. If $f=\begin{pmatrix} a_1 & a_2 & a_3 \dots a_n \\ b_1 & b_2 & b_3 \dots b_n \end{pmatrix}$ be a permutation of degree n, then the inverse of f *i.e.*, f^{-1} is obtained by interchanging the rows of f. Thus $f^{-1}=\begin{pmatrix} b_1 & b_2 & b_3 \dots b_n \\ a_1 & a_2 & a_3 \dots a_n \end{pmatrix}$.

Note 2. If we are to write the set of all permutations of degree n, it is immaterial whatever symbols we use to denote the elements of the set S. We can use the numbers 1, 2, 3,..., n or we can use the letters $a_1, a_2, \dots, a_n$ or any symbols.

Note 3. The group P_n of all permutations of degree n is called *the symmetric group* of degree n or *the symmetric group of order n !*

§ 14. **Cyclic Permutations. Definition.** *Suppose f is a permutation of degree n on a set S having n distinct elements. Let it be possible to arrange m elements of the set S in a row in such a way that the f-image of each element in the row is the element which follows it, the f-image of the last element is the first element and the remaining n - m elements of the set S are left unchanged by f. Then f is called a cyclic permutation or a cycle of length m or an m-cycle.*

(Rajasthan 1976)

By the length of a cycle we mean the number of objects permuted by the cycle.

For example the permutation $f=\begin{pmatrix}1 & 2 & 5 & 3 & 6 & 4\\ 2 & 4 & 5 & 1 & 6 & 3\end{pmatrix}$ is cyclic. It can be represented by the cycle (1 2 4 3) which means that each element in the bracket is replaced by the element following it, the last element is replaced by the first element and the missing elements 5 and 6 are left unchanged. Thus 1 is replaced by 2, 2 by 4, 4 by 3, 3 by 1, 5 by 5 and 6 by 6. The length of this cycle is 4.

Similarly the permutation $g=\begin{pmatrix}1 & 2 & 3 & 5 & 4 & 6\\ 2 & 3 & 4 & 6 & 5 & 1\end{pmatrix}$ is cyclic. We can write it in the form of the cycle (1 2 3 4 5 6) which is of length 6.

However the permutation $h=\begin{pmatrix}1 & 2 & 4 & 3 & 5 & 6\\ 2 & 1 & 3 & 4 & 5 & 6\end{pmatrix}$ is not cyclic.

Permutations represented by a cycle. (1 3 4 2 6) is a cycle of length 5. Suppose it represents a permutation of degree 9 on a set S consisting of the elements 1, 2,..., 9. Then the permutation represented will be

$$\begin{pmatrix}1 & 3 & 4 & 2 & 6 & 5 & 7 & 8 & 9\\ 3 & 4 & 2 & 6 & 1 & 5 & 7 & 8 & 9\end{pmatrix}$$

i.e., the image of each element in the cycle (1 3 4 2 6) is the element which follows it, the image of the last element 6 is the first element 1 and the missing elements 5, 7, 8, 9 are their images themselves.

However if the cycle (1 3 4 2 6) represents a permutation of degree 6 on six symbols 1, 2, 3, 4, 5, 6 then the corresponding permutation will be

$$\begin{pmatrix}1 & 3 & 4 & 2 & 6 & 5\\ 3 & 4 & 2 & 6 & 1 & 5\end{pmatrix}.$$

Important Note. *A cycle does not change by changing the places of its elements provided their cyclic order is not changed.*

Thus (1 2 3 4)=(2 3 4 1)=(3 4 1 2)=(4 1 2 3).

Also (1 2)=(2 1), (2 3)=(3 2).

Transpositions. Definition. **(Raj. 1978)**

A cycle of length two is called a transposition. Thus the cycle (1 3) is a transposition. It will represent a permutation in which the image of 1 is 3, the image of 3 is 1 and the remaining missing elements are left unchanged.

If the transposition (2 3) is a permutation of degree 3 on three symbols 1, 2, 3 then the corresponding permutation will be

$$\begin{pmatrix}1 & 2 & 3\\ 1 & 3 & 2\end{pmatrix}.$$

A cycle of length one means that the image of the element involved is the element itself and the missing elements are left unchanged. Thus all the elements are left unchanged. Therefore every cycle of length one will represent the identity permutation.

Multiplication of Cycles. We multiply cycles by multiplying the permutations represented by them. For example if the cycles (1 2 3) and (5 6 4 1) represent permutations of degree 6 on six symbols 1, 2, 3, 4, 5, 6, then

$$(1\ 2\ 3)\ (5\ 6\ 4\ 1)=\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6\\ 2 & 3 & 1 & 4 & 5 & 6\end{pmatrix}\begin{pmatrix}5 & 6 & 4 & 1 & 2 & 3\\ 6 & 4 & 1 & 5 & 2 & 3\end{pmatrix}$$
$$=\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6\\ 2 & 3 & 5 & 1 & 6 & 4\end{pmatrix}=(1\ 2\ 3\ 5\ 6\ 4).$$

Since a cycle of length one represents the identity permutation, therefore (1) (2 3 4) (6) = (2 3 4).

Disjoint Cycles. *Two cycles are said to be disjoint if they have no symbols in common.* For example (1 3 5) and (2 6 8 9) are disjoint cycles while (1 3 4) and (2 3 5 6) are not disjoint.

Theorem. *If f and g are two disjoint cycles, then $fg=gf$ i.e., the product of disjoint cycles is commutative.*

Proof. The cycles f and g have no symbols common. Therefore the elements permuted by f are left unchanged by g and also the elements permuted by g remain the same under f. Therefore we shall have $fg=gf$.

Now we shall give an example to illustrate this theorem. Let $f=(1\ 2\ 3)$ and $g=(4\ 5)$ represent two permutations on 5 symbols 1, 2,..., 5.

$$\text{Then } fg=\begin{pmatrix}1 & 2 & 3 & 4 & 5\\ 2 & 3 & 1 & 4 & 5\end{pmatrix}\begin{pmatrix}1 & 2 & 3 & 4 & 5\\ 1 & 2 & 3 & 5 & 4\end{pmatrix}$$
$$=\begin{pmatrix}1 & 2 & 3 & 4 & 5\\ 2 & 3 & 1 & 5 & 4\end{pmatrix}=\begin{pmatrix}4 & 5 & 1 & 2 & 3\\ 5 & 4 & 2 & 3 & 1\end{pmatrix}$$
$$=\begin{pmatrix}4 & 5 & 1 & 2 & 3\\ 5 & 4 & 1 & 2 & 3\end{pmatrix}\begin{pmatrix}1 & 2 & 3 & 4 & 5\\ 2 & 3 & 1 & 4 & 5\end{pmatrix}$$
$$=(4\ 5)\ (1\ 2\ 3)=gf.$$

Inverse of a cyclic permutation. *To prove that* $(1\ 2\ 3 \ldots n)^{-1}=(n\ n-1 \ldots .3\ 2\ 1)$ *i.e., to write the inverse of a cycle we should write its elements in the reverse order.*

Proof. We have $(1\ 2\ 3 \ldots n)\ (n \cdots 3\ 2\ 1)$

$$=\begin{pmatrix}1 & 2 & 3 \ldots n-1 & n\\ 2 & 3 & 4 \ldots n & 1\end{pmatrix}\begin{pmatrix}n & \ldots 4 & 3 & 2 & 1\\ n-1 & \ldots 3 & 2 & 1 & n\end{pmatrix}$$

$$=\begin{pmatrix}1 & 2 & 3\ldots n-1 & n\\ 1 & 2 & 3\ldots n-1 & n\end{pmatrix}=I.$$

Also $$(n\ldots 3\ \ 2\ \ 1)(1\ \ 2\ \ 3\ldots n)=I.$$

$$\therefore\ (1\ \ 2\ \ 3\ldots n)^{-1}=(n\ldots 3\ \ 2\ \ 1).$$

In particular *every transposition is its own inverse.* If $(1\ \ 2)$ is a transposition, then $(1\ \ 2)^{-1}=(2\ \ 1)=(1\ \ 2)$.

Inverse of a product of cyclic permutations. If f and g are any two cycles, then we have

$$(fg)^{-1}=g^{-1}f^{-1}.\quad \text{Also } (fgh)^{-1}=h^{-1}g^{-1}f^{-1}.$$

If f and g are disjoint cycles then $(fg)^{-1}=(gf)^{-1}=f^{-1}g^{-1}$.

Thus $[(1\ 2\ 3)(4\ 5)(2\ 6)]^{-1}=(2\ 6)^{-1}(4\ 5)^{-1}(1\ 2\ 3)^{-1}$

$=(6\ 2)(5\ 4)(3\ 2\ 1)$.

Also $[(1\ 3\ 5)(2\ 4)]^{-1}=(1\ 3\ 5)^{-1}(2\ 4)^{-1}=(5\ 3\ 1)(4\ 2)$.

We shall now give some important results on the product of permutations.

Theorem 1. *Every permutation can be expressed as a product of disjoint cycles.* **(Madras 1974)**

Verification. Let $f=\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9\\ 2 & 3 & 1 & 4 & 8 & 6 & 9 & 7 & 5\end{pmatrix}$ be a permutation of degree 9 on the set $\{1, 2,\ldots, 9\}$.

We have $f=(4)(6)(1\ 2\ 3)(5\ 8\ 7\ 9)$.

Explanation. First we put down cycles of length one with the help of elements which remain unchanged under f.

Then we start with an element which is not left unchanged. Thus we start with 1. After 1 we write the image of 1 which is 2. After 2 we write the image of 2 which is 3. After 3 we do not write any element but we close the bracket since the image of the last element 3 is the first element 1 of the bracket.

Now we start a new bracket. In this bracket we write an element which has not yet been written. Thus we write 5. After 5 we write the image of 5 which is 8. After 8 we write the image of 8 which is 7. After 7 we write the image of 7 which is 9. After 9 we close the bracket since the image of 9 is the first element 5.

Now our work is finished since each element has been included in one or the other bracket.

Since cycles of length one represent identity permutation, therefore we can omit them. Thus we can write

$$f=(1\quad 2\quad 3)(5\quad 8\quad 7\quad 9).$$

Also we can write $f=(5\ 8\ 7\ 9)\ (1\ 2\ 3)$ because product of disjoint cycles is commutative.

Theorem 2. *Every cycle can be expressed as a product of transpositions in infinitely many ways.*

Verification. Consider the cycle $(1\ 2\ 3 \ldots n)$ of length n. By actual calculation we see that

$$(1\ 2\ 3 \ldots n)=(1\ 2)\ (1\ 3)\ (1\ 4) \ldots (1\ n-1)\ (1\ n).$$

More generally the n-cycle

$$(a_1\ a_2\ a_3 \ldots a_n)=(a_1\ a_2)\ (a_1\ a_3) \ldots (a_1\ a_n).$$

The 4-cycle $(2\ 3\ 5\ 4)=(2\ 3)\ (2\ 5)\ (2\ 4)$.

Since $(2\ 3\ 5\ 4)=(3\ 5\ 4\ 2)$, therefore we can write

$$(2\ 3\ 5\ 4)=(3\ 5)\ (3\ 4)\ (3\ 2).$$

Also $(1\ 2)\ (2\ 1)=$ Identity permutation.

$\therefore$ We can write $(2\ 3\ 5\ 4)=(2\ 3)\ (1\ 2)\ (2\ 1)\ (2\ 5)\ (2\ 4)$.

Again $(2\ 5\ 3)\ (3\ 5\ 2)$ is also identity permutation.

$\therefore$ We can write $(2\ 3\ 5\ 4)=(2\ 3\ 5\ 4)\ (2\ 5\ 3)\ (3\ 5\ 2)$

$$= (2\ 3)\ (2\ 5)\ (2\ 4)\ (2\ 5)\ (2\ 3)\ (3\ 5)\ (3\ 2).$$

Now consider the 5-cycle $(1\ 2\ 3\ 4\ 5)$.

We have $(1\ 2\ 3\ 4\ 5)=(1\ 2)\ (1\ 3)\ (1\ 4)\ (1\ 5)$

$$=(1\ 2)\ (2\ 4)\ (4\ 2)\ (1\ 3)\ (1\ 2)\ (2\ 1)\ (1\ 4)\ (1\ 5).$$

Thus every cycle can be expressed as a product of transpositions in infinitely many ways. But we see that in the case of any cycle the number of transpositions is either always odd or always even.

Theorem 3. *Every permutation can be expressed as a product of transpositions in infinite many ways.*

Combining together the results of Theorem 1 and Theorem 2, we immediately get the result of this theorem.

§ 15. Even and Odd Permutations. Definition. *A permutation is said to be an even permutation if it can be expressed as a product of an even number of transpositions; otherwise it is said to be an odd permutation.* **(Banaras 1971; Rajasthan 77, 78)**

This definition will be meaningless if a permutation can be expressed sometimes as a product of an odd number of transpositions and sometimes as a product of an even number of transpositions. So we have the following theorem.

Theorem. *A permutation cannot be both even and odd i.e., if a permutation f is expressed as a product of transpositions then the number of transpositions is either always even or always odd.*

(Banaras 1968)

Proof. Consider the polynomial P in n distinct symbols $x_1, x_2, \ldots, x_n$ defined as the product of all factors of the type x_i-x_j where $i<j$.

$$\text{Thus}\quad P=\prod_{i<j=1}^{n}(x_i-x_j)$$

$$=(x_1-x_2)(x_1-x_3)(x_1-x_4)\ldots(x_1-x_n)$$
$$(x_2-x_3)(x_2-x_4)\ldots(x_2-x_n)$$

$$(x_{n-2}-x_{n-1})(x_{n-2}-x_n)(x_{n-1}-x_n).$$

Suppose now f is any permutation on n symbols $1, 2, 3, \ldots, n$.

By fP we mean the polynomial obtained by changing in P the subscripts $1, 2, \ldots, n$ of x's as prescribed by f.

[For example, taking $n=4$, we have

$P=(x_1-x_2)(x_1-x_3)(x_1-x_4)(x_2-x_3)(x_2-x_4)(x_3-x_4)$.

If $f=(1\ 2\ 3)$, then we have

$fP=(x_2-x_3)(x_2-x_1)(x_2-x_4)(x_3-x_1)(x_3-x_4)(x_1-x_4)$.

In particular if $f=$the transposition $(2\ 3)$, we have

$fP=(x_1-x_3)(x_1-x_2)(x_1-x_4)(x_3-x_2)(x_3-x_4)(x_2-x_4)$

$=-P$ *i.e.*, the effect of a transposition on P is to change the sign of P].

Now the operation by a transposition (r, s), where $r<s$ has the following effect on P.

(i) Any factor of P which contains neither the suffix r nor s remains unchanged.

(ii) The single factor (x_r-x_s) changes its sign on replacing r by s and s by r.

(iii) The remaining factors which contain either the suffix r or s but not both can be grouped into the following three types of products :

(a) $[(x_1-x_r)(x_1-x_s)]\,[(x_2-x_r)(x_2-x_s)]\ldots$
$[(x_{r-1}-x_r)(x_{r-1}-x_s)]$

(b) $[(x_r-x_{r+1})(x_{r+1}-x_s)]\,[(x_r-x_{r+2})(x_{r+2}-x_s)]\ldots$
$[(x_r-x_{s-1})(x_{s-1}-x_s)]$

(c) $[(x_r-x_{s+1})(x_s-x_{s+1})]\,[(x_r-x_{s+2})(x_s-x_{s+2})]\ldots$
$[(x_r-x_n)(x_s-x_n)]$.

On replacing r by s and s by r, the signs of all these types of products do not change.

Hence the net effect of the transposition (r, s) on P is to change the sign of P *i.e.*, P operated upon by a transposition becomes $-P$.

Therefore if the permutation f can be expressed as the product of p transpositions then $fP=(-1)^p P$ and if f can be expressed as the product of q transpositions then

$$fP=(-1)^q P.$$

This gives $(-1)^p P=(-1)^q P$

$\Rightarrow (-1)^p=(-1)^q$

$\Rightarrow$ either p and q are both even or both odd.

Hence the theorem.

Cor. 1. *A cycle of length n can be expressed as the product of $n-1$ transpositions. Therefore a cycle of length n will be an even permutation if n is odd and it will be an odd permutation if n is even.*

In particular every transposition is an odd permutation.

Cor. 2. *Identity permutation is always an even permutation.*

If I is the identity permutation then I can be expressed as the product of two transpositions. For example we can write

$$I=(1\ 2)\ (2\ 1).$$

$\therefore$ I is an even permutation.

Cor. 3 *The product of two even permutations is an even permutation* Suppose f and g are two even permutations. Further suppose that f can be expressed as the product of r transpositions and g can be expressed as the product of s transpositions. Then r and s are both even. Now fg can be expressed as the product of $r+s$ transpositions. Since $r+s$ is even, therefore fg is an even permutation.

Cor. 4. *The product of two odd permutations is an even permutation.* Proceed as in Cor. 3. If r and s are both odd then also $r+s$ is even.

Cor. 5. *The product of an even permutation and an odd permutation is an odd permutation. Similarly the product of an odd permutation and an even permutation is an odd permutation.*

Cor. 6. *The inverse of an even permutation is an even permutation and the inverse of an odd permutation is an odd permutation.*

Suppose f is an even permutation. If f^{-1} is the inverse of f, then $f^{-1}f=I$ (identity permutation).

Now I is an even permutation and f is also an even permutation. Therefore f^{-1} cannot be an odd permutation otherwise

$f^{-1}.f$ will be an odd permutation. Hence f^{-1} is an even permutation.

Similarly if f is an odd permutation, then f^{-1} must also be an odd permutation.

Total number of even permutations of degree n.

Theorem. *Of the n ! permutations on n symbols, ½n ! are even permutations and ½n ! are odd permutations.*

(Meerut 1968; Madurai 78; Banaras 71)

Proof. Out of the n ! permutations on n symbols let the even permutations be $e_1, e_2, \ldots, e_m$ and the odd permutations be $o_1, o_2, \ldots, o_k$.

Since a permutation is either an even permutation or an odd permutation but not both, therefore $m+k=n$!.

If P_n be the set of all permutations of degree n, then

$$P_n=\{e_1, e_2, \ldots, e_m, o_1, o_2, \ldots, o_k\}.$$

Let $t \in P_n$ and suppose t is a transposition.

Since P_n is a group with respect to permutation multiplication, therefore $te_1, te_2, \ldots, te_m, to_1, to_2, \ldots, to_k$ are all elements of P_n. Obviously $te_1, te_2, \ldots, te_m$ are all odd permutations and $to_1, to_2, \ldots, to_k$ are all even permutations.

Now no two of the permutations $te_1, te_2, \ldots, te_m$ are equal because

$te_i=te_j \Rightarrow e_i=e_j$ (by left cancellation law in the group P_n).

Therefore if $e_i \neq e_j$, then $te_i \neq te_j$.

Thus the m odd permutations $te_1, \ldots, te_m$ are distinct elements of P_n. But we have supposed that P_n contains exactly k odd permutations. Therefore m cannot be greater than k. Thus

$$m \leqslant k. \qquad \ldots(1)$$

Similarly, we can show that the k even permutations $to_1, to_2, \ldots, to_k$ are distinct elements of P_n. Therefore, we must have

$$k \leqslant m. \qquad \ldots(2)$$

From (1) and (2), it follows that $m=k=\dfrac{n\,!}{2}$.

Note. If A_n is the set of all even permutations of degree n then $A_n \subset P_n$ and A_n contains $\dfrac{n\,!}{2}$ elements. *The set A_n is called an Alternating set of permutations of degree n.*

Group of all even permutations of degree n.

Theorem. *The set A_n of all even permutations of degree n* *forms a finite group of order* $\frac{n!}{2}$ *with respect to permutation multiplication.* **(Gorakhpur 1970; Banaras 70; Lucknow 69)**

Proof. The product of two even permutations is also an even permutation. Therefore the set A_n is closed with respect to multiplication of permutations as composition.

We know that multiplication of permutations is an associative composition.

If I is the identity permutation of degree n then I is an even permutation. Therefore $I \in A_n$. Now we have

$$If=f=fI \; \forall \, f \in A_n.$$

$\therefore$ I is the identity element.

Let f be any even permutation of degree n. If f^{-1} is the inverse of f in the group of all permutations of degree n, then f^{-1} is also an even permutation because $f^{-1}f=I$ (an even permutation).

Thus $f \in A_n \Rightarrow$ that there exists $f^{-1} \in A_n$ such that

$$f^{-1}f=I=ff^{-1}.$$

$\therefore$ every element of A_n possesses inverse.

The total number of all even permutations of degree n is $\frac{n!}{2}$. Thus there are $\frac{n!}{2}$ elements in the set A_n.

$\therefore$ A_n forms a finite group of order $\frac{n!}{2}$ with respect to multiplication of permutations.

Note. The product of two odd permutations is an even permutation. Therefore the set of all odd permutations is not closed with respect to multiplication. Therefore it will not be a group.

Ex. 1. *Write the following permutations as the product of disjoint cycles.*

(*a*) $f=\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 3 & 4 & 5 & 1 & 6 & 7 & 9 & 8 \end{pmatrix}$

(*b*) $g=\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 3 & 1 & 2 \end{pmatrix}$

Solution. (a) We have

$$f=(6)\,(7)\,(1\ 2\ 3\ 4\ 5)\,(8\ 9)$$

or $f=(1\ 2\ 3\ 4\ 5)\,(8\ 9)$, omitting cycles of length 1 as they represent identity permutation.

(b) We have $g=(1\ 6\ 2\ 5)\,(3\ 4)$.

Ex. 2. *Express the following permutations as the product of disjoint cycles :*

(*a*) $f=(1\ 2\ 3)\ (4\ 5)\ (1\ 6\ 7\ 8\ 9)\ (1\ 5)$,

(*b*) $g=(1\ 2)\ (1\ 2\ 3)\ (1\ 2)$,

(*c*) $h=(1\ 3\ 2\ 5)\ (1\ 4\ 3)\ (2\ 5\ 1)$.

Solution. (a) We shall start with the element 1. Under the first cycle 1 goes to 2, under the second cycle 2 goes to 2, under the third cycle 2 goes to 2 and under the fourth cycle 2 goes to 2. Therefore under the permutation f, 1 goes to 2. Now we are to find the image of 2 under f. In the first cycle 2 goes to 3, in each of the other three cycles 3 is left unchanged. Therefore under f, 2 goes to 3. Now we are to find the image of 3 under f. In the first cycle 3 goes to 1, in the second 1 goes to 1, in the third 1 goes to 6 and in the fourth 6 goes to 6. Thus 3 goes to 6 under f. Now the image of 6 under f is 7. The image of 7 under f is 8, the image of 8 under f is 9. Now in the first two cycles 9 is left unchanged. In the third cycle 9 goes to 1 and in the fourth 1 goes to 5. Therefore 9 goes to 5 under f.

Now under f the element 5 goes to 4 and 4 goes to 1.

Since we started with the element 1, therefore we stop at this stage and thus we have obtained the first cycle as

$$(1\ 2\ 3\ 6\ 7\ 8\ 9\ 5\ 4).$$

Since this single cycle contains all the symbols, therefore f is a cyclic permutation and we have

$$f=(1\ 2\ 3\ 6\ 7\ 8\ 9\ 5\ 4).$$

(b) We have $g=(1\ 3\ 2)$.

(c) We have $h=(1\ 2)\ (3\ 5\ 4)$.

Ex. 3. *Determine which of the following are even permutations :*

(*a*) $f=(1\ 2\ 3)\ (1\ 2)$. (*b*) $g=(1\ 2\ 3\ 4\ 5)\ (1\ 2\ 3)\ (4\ 5)$

(*c*) $h=(1\ 2)\ (1\ 3)\ (1\ 4)\ (2\ 5)$.

Solution. (a) We can write $f=(1\ 2)\ (1\ 3)\ (1\ 2)$. The number of transpositions is 3 *i.e.*, odd. Therefore f is an odd permutation.

(b) We can write $g=(1\ 2)\ (1\ 3)\ (1\ 4)\ (1\ 5)\ (1\ 2)\ (1\ 3)\ (4\ 5)$.

The number of transpositions is 7 *i.e.*, odd.

Therefore g is an odd permutation.

(c) $h=(1\ 2)\ (1\ 3)\ (1\ 4)\ (2\ 5)$. The number of transpositions is 4 *i.e.*, even. Therefore h is an even permutation.

Ex. 4. *Show that the set P_3 of all permutations on three symbols* 1, 2, 3 *is a finite non-abelian group of order* 6 *with respect to permutation multiplication as composition.*

(Rajasthan 1977; Meerut 72)

Solution. We have $P_3=\{I, (1\ 2), (2\ 3), (3\ 1), (1\ 2\ 3), (1\ 3\ 2)\}$ where I is the identity permutation. There are 6 elements in the set P_3. Let $I=f_1$, $(1\ 2)=f_2$, $(2\ 3)=f_3$, $(3\ 1)=f_4$, $(1\ 2\ 3)=f_5$, $(1\ 3\ 2)=f_6$. We now prepare a composition table for P_3.

Product of permutations	f_1	f_2	f_3	f_4	f_5	f_6
f_1	f_1	f_2	f_3	f_4	f_5	f_6
f_2	f_2	f_1	f_6	f_5	f_4	f_3
f_3	f_3	f_5	f_1	f_6	f_2	f_4
f_4	f_4	f_6	f_5	f_1	f_3	f_2
f_5	f_5	f_3	f_4	f_2	f_6	f_1
f_6	f_6	f_4	f_2	f_3	f_1	f_5

For preparing this composition table we have actually multiplied the permutations. Thus

$f_2 f_3=(1\ 2)\ (2\ 3)=(1\ 3\ 2)=f_6$,
$f_2 f_4=(1\ 2)\ (3\ 1)=(1\ 2\ 3)=f_5$,
$f_3 f_3=(2\ 3)\ (2\ 3)=$ identity permutation $=f_1$,
$f_5 f_5=(1\ 2\ 3)\ (1\ 2\ 3)=(1\ 3\ 2)=f_6$, and so on.

Since all the entries in the table are elements of P_3, therefore P_3 is closed with respect to multiplication of permutations.

Multiplication of permutations is an associative composition.

The identity permutation f_1 is the identity element since $f_1 f_1=f_1$, $f_1 f_2=f_2=f_2 f_1$, $f_1 f_3=f_3=f_3 f_1$ and so on.

Every element possesses inverse. The inverse of f_1 is f_1, the inverse of f_2 is f_2, the inverse of f_3 is f_3, the inverse of f_4 is f_4, the inverse of f_5 is f_6 and the inverse of f_6 is f_5. [Note that $f_5 f_6=f_1=f_6 f_5$, therefore $f_5^{-1}=f_6$ and $f_6^{-1}=f_5$].

The composition is not commutative since $f_2 f_3=f_6$ and $f_3 f_2=f_5$. Thus $f_2 f_3 \neq f_3 f_2$.

$\therefore$ P_3 is a finite non-abelian group of order 6 with respect to permutation multiplication.

Ex. 5. *Write down all the permutations on four symbols* 1, 2 3, 4. *Which of these permutations are even ?*

Solution. There will be 4 ! *i.e.*, 24 permutations of degree 4. If P_4 is the set of all these permutations, then

$P_4=\{(1), (12), (13), (14), (23), (24), (34), (123), (132), (124), (142), (134), (143), (234), (243), (12)(34), (23)(14), (31)(24), (1234), (1243), (1324), (1342), (1423), (1432)\}$.

If A_4 is the set of all even permutations of degree 4, then A_4 will have $\frac{1}{2}\times 4$! *i.e.*, 12 elements.

Thus $A_4=\{(1), (123), (132), (124), (142), (134), (143), (234), (243), (12)(34), (23)(14), (31)(24)\}$.

Ex. 6. *Show that the four permutations* I, (ab), (cd), $(ab)(cd)$ *on four symbols* a, b, c, d *form a finite abelian group with respect to the permutation multiplication.*

Solution. Let $I=f_1$, $(ab)=f_2$, $(cd)=f_3$, $(ab)(cd)=f_4$. To prepare the composition table, we observe that f_2 and f_3 are transpositions. Therefore $f_2f_2=f_1$, $f_3f_3=f_1$. Also f_2 and f_3 are disjoint cycles. Therefore $f_2f_3=f_3f_2=f_4$.

Further $f_2f_4=(ab)(ab)(cd)=I(cd)=(cd)=f_3$.

Similarly $f_3f_4=(cd)(ab)(cd)=(cd)(cd)(ab)=I(ab)=(ab)=f_2$

$[\because (ab)(cd)=(cd)(ab)]$.

Also $f_4f_4=(ab)(cd)(ab)(cd)$

$=(ab)(ab)(cd)(cd)=(I)(I)=I=f_1$.

Similarly making all other calculations, the composition table is

Product of permutations	f_1	f_2	f_3	f_4
f_1	1	f_2	f_3	f_4
f_2	f_2	f_1	f_4	f_3
f_3	f_3	f_4	f_1	f_2
f_4	f_4	f_3	f_2	f_1

From the table it is clear that

(i) All the entries in the composition table are elements of the given set. Therefore the closure axiom is satisfied.

(ii) f_1 is the identity element.

(iii) Each element possesses inverse. In fact $f_1^{-1}=f_1$, $f_2^{-1}=f_2$, $f_3^{-1}=f_3$, $f_4^{-1}=f_4$.

(iv) The composition is commutative.

Further the multiplication of permutations is an associative composition.

$\therefore$ the given set is a finite abelian group of order 4 with respect to the permutation multiplication.

Ex. 7. *Show that the eight permutations* (a), $(abcd)$, $(ac)\,(bd)$, $(adcb)$, $(ab)\,(cd)$, $(bc)\,(ad)$, (bd), (ac) *on four symbols* a, b, c, d *form a finite non-abelian group with respect to permutation multiplication.*

Solution. Proceed as in Ex. 6.

Exercises

1. Write down all the permutations on three symbols a, b, c. Which of these permutations are even ? **(Rajasthan 1977)**

Ans. I (identity permutation), (ab), (bc), (ca), (abc), (acb); even permutations are I, (abc), (acb).

2. Find the orders of the groups :

(i) the group S_4 (symmetric group), (ii) the group A_4 (alternating group). **(Meerut 1976)**

Ans. (i) 24, (ii) 12.

3. Define a permutation. If $A=\begin{pmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{pmatrix}$ and $B=\begin{pmatrix}1 & 2 & 3\\ 3 & 1 & 2\end{pmatrix}$, find AB and BA. **(Meerut 1973, 75)**

Ans. Both AB and BA are identity permutations.

4. Find the inverse of each of the following permutations :

(i) $\begin{pmatrix}1 & 2 & 3 & 4\\ 1 & 3 & 4 & 2\end{pmatrix}$, (Luck. 1980) (ii) $\begin{pmatrix}1 & 2 & 3 & 4\\ 3 & 4 & 1 & 2\end{pmatrix}$,

(iii) $\begin{pmatrix}1 & 2 & 3 & 4 & 5\\ 2 & 3 & 1 & 5 & 4\end{pmatrix}$. **(Meerut 1976)**

Ans. (i) $\begin{pmatrix}1 & 2 & 3 & 4\\ 1 & 4 & 2 & 3\end{pmatrix}$, (ii) $\begin{pmatrix}1 & 2 & 3 & 4\\ 3 & 4 & 1 & 2\end{pmatrix}$,

(iii) $\begin{pmatrix}1 & 2 & 3 & 4 & 5\\ 3 & 1 & 2 & 5 & 4\end{pmatrix}$.

5. Decompose the following permutations into transpositions :—

(i) $\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7\\ 6 & 5 & 2 & 4 & 3 & 1 & 7\end{pmatrix}$ **(Meerut 1976)**

(ii) $\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7 & 8\\ 3 & 1 & 4 & 7 & 2 & 5 & 8 & 6\end{pmatrix}$. **(Meerut 1977, 80)**

Ans. (i) $(1\ 6)\,(2\ 5)\,(2\ 3)$;

(ii) $(1\ 3)\,(1\ 4)\,(1\ 7)\,(1\ 8)\,(1\ 6)\,(1\ 5)\,(1\ 2)$.

6. Examine whether the following permutation is even or odd :

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 5 & 4 & 3 & 6 & 1 & 7 & 9 & 8 \end{pmatrix}.$$ **(Madras 1974)**

Ans. Odd.

7. Given that $f=(1\ 3\ 2\ 5)\ (1\ 4\ 3)\ (2\ 5\ 1)$ is a permutation on five symbols. Express it as a product of disjoint cycles. Also find the inverse of f and express it as a product of disjoint cycles.

Ans. $f=(1\ 2)\ (3\ 5\ 4)$; $f^{-1}=(1\ 2)\ (3\ 4\ 5)$.

8. Prove that a cycle containing an odd number of symbols is an even permutation whereas a cycle containing an even number of symbols is an odd permutation.

9. Show that the set G of four permutations I, $(1\ 2)\ (3\ 4)$, $(1\ 3)\ (2\ 4)$, and $(1\ 4)\ (2\ 3)$ on four symbols 1, 2, 3, 4 is an abelian group with respect to the permutation multiplication.

(This group is known as the four group V_4 of Klein).

10. Show that the n ! permutations of n objects form a group with respect to permutation multiplication. Verify this theorem by taking the set of all permutations of elements 1, 2, 3.

(Meerut 1973)

11. Prove that the set A_3 of three permutations

$$(a),\ (a\ b\ c),\ (a\ c\ b)$$

on three symbols a, b. c forms a finite abelian group with respect to the permutation multiplication.

12. Show that if S has more than two elements, then the symmetric group S_n is not abelian.

[**Hint.** Refer Example 6 page 64].

13. Show that the group S_3 is non-abelian. **(Meerut 1977)**

14. Give an example of a finite non-abelian group.

(Meerut 1978 ; G. N.D U. Amritsar 82)

15. G is a group and a is a fixed element of G. Show that the mapping $f_a : G\rightarrow G$ defined by $f_a\ (x)=ax\ \forall x\in G$ is a permutation of G *i.e.*, is a one-to-one mapping of G onto G. **(Marathwada 1972)**

§ 16. Integral powers of an element of a group (composition denoted multiplicatively).

Suppose G is a group and the composition has been denoted multiplicatively. Let $a \in G$. Then by closure property a, aa, aaa, $aaaa$, etc. are all elements of G. Since the composition in G obeys general associative law, therefore $aaa\ .\ a$ to n factors is independent of the manner in which the factors may be grouped.

If n is a positive integer, we define $a^n = aaa \ldots a$ to n factors. Obviously $a^n \in G$. In particular $a^1 = a$, $a^2 = aa$, $a^3 = aaa = a^2 a$ and so on.

If e is the identity element of the group G, then we define

$$a^0 = e.$$

If n is a positive integer then $-n$ is a negative integer. Now we define $a^{-n} = (a^n)^{-1}$ where $(a^n)^{-1}$ is the inverse of a^n in G. Thus $a^{-n} \in G$.

Thus we have defined a^n for all integral values of n positive, zero or negative.

However, if n is a positive integer, then according to our definition

$$a^{-n} = (a^n)^{-1} = (aaa \ldots a \text{ upto } n \text{ factors})^{-1}$$
$$= (a^{-1})\,(a^{-1})\,(a^{-1}) \ldots (a^{-1}) \text{ upto } n \text{ factors} = (a^{-1})^n.$$

Thus we are free to write $(a^{-1})^n$ or $(a^n)^{-1}$ in place of a^{-n}.

In particular $a^{-4} = (a^4)^{-1} = (a^{-1})^4$.

Integral multiples of an element of a group. Composition denoted additively.

If in a group G the composition has been denoted additively, then in place of using the word integral powers of an element of a group we use the word integral multiples of an element of a group. The difference is only of notation otherwise the meaning is the same. Thus in this case if n is a positive integer we write na in place of a^n and we define $na = a + a + \ldots + a$ upto n terms.

In place of a^0 we write $0a$. Thus we define $0a = e$ where e is the identity of G.

If n is a positive integer, then in place of a^{-n} we write $(-n)a$.

Thus we define $(-n)\,a = -(na)$, where $-(na)$ denotes the inverse of na in G. Also we can prove that $-(na) = n\,(-a)$.

Thus we can write $-(na)$ or $n\,(-a)$ in place of $(-n)\,a$.

In particular $6a = a + a + \ldots + a$ upto 6 terms.

$$(-6)\,a = -(6a) \text{ i.e., } -(a + a + \ldots \text{upto 6 terms})$$

or $$(-6)\,a = 6\,(-a) \text{ i.e., } (-a) + (-a) + \ldots \text{upto 6 terms}.$$

In multiplicative notation the following laws of indices can be easily proved :

$$a^m a^n = a^{m+n},$$
and $$(a^m)^n = a^{mn},$$

$\forall\, a \in G$ and $\forall\, m, n \in \mathbf{I}$ where $\mathbf{I}$ is the set of integers.

In additive notation the following laws of multiples can be easily proved :

$$ma+na=(m+n)\,a,$$
$$n\,(ma)=(nm)\,a,$$
$$\forall\ a \in G \text{ and } \forall\ m, n \in \mathbf{I}.$$

§ 17. Order of an element of a group. Definition.

(Guru Nanak 1982 ; Banaras 70 ; Kolhapur 73 ; Rajasthan 77; Meerut 79; Lucknow 80; Gorakhpur 70)

Suppose G is a group and the composition has been denoted multiplicatively. By the order of an element $a \in G$ is meant the least positive integer n, if one exists, such that

$$a^n=e \text{ (the identity of } G).$$

If there exists no positive integer n such that $a^n=e$, then we say that a is of **infinite order** *or of* **zero order.**

We shall use the symbol $o\,(a)$ to denote the order of a.

In additive notation we use the words $na=e$ in place of $a^n=e$.

Important. If there exists a positive integer m such that $a^m=e$, then the order of a is definitely finite. Also we must have $o(a) \leqslant m$. When $a^m=e$, then the question of order of a being greater than m does not arise. At the most it can be equal to m. If m itself is the least positive integer such that $a^m=e$, then we will have

$$o\,(a)=m.$$

In any group the identity element e is always of order one and it is the only element of order one.

We have $e^1=e \Rightarrow o\,(e)=1$. Also if $o\,(a)=1$, then $a^1=a=e$.

Example 1. *Let us find the order of each element of the multiplicative group* $\{1, -1, i, -i\}$. **(Lucknow 1980)**

Since 1 is the identity element, therefore $o\,(1)=1$.

Now $(-1)^1=-1$, $(-1)^2=(-1)(-1)=1$ (*i.e.*, identity element).

$\therefore\ o\,(-1)=2.$

Again $(i)^1=i$, $(i)^2=-1$, $i^3=-i$, $i^4=1$ (*i.e.*, identity element).

$\therefore\ o\,(i)=4.$

Finally $(-i)^1=-i$, $(-i)^2=-1$, $(-i)^3=i$, $(-i)^4=1$.

$\therefore\ o\,(-i)=4.$

Example 2. *Let us find the orders of some elements of the group* $\{0, 1, 2, 3, 4, 5\}$, *the composition being addition modulo* 6.

Since 0 is the identity element, therefore $o\,(0)=1$.

Now $(1)^1=1$, $1^2=1+_6 1=2$, $1^3=1+_6 1+_6 1=1+_6 2=3$, $1^4=1+_6 1+_6 1+_6 1=1+_6 3=4$, $1^5=1+_6 1^4$ $=1+_6 4=5$, $1^6=1+_6 1^5=1+_6 5=0$ (*i.e.*, identity element).

$\therefore\ o\,(1)=6.$

Again $2^1=2$, $2^2=2+_6 2=4$, $2^3=2+_6 2^2=2+_6 4=0$ (identity element).

$$\therefore \quad o(2)=3.$$

Further $3^1=3$, $3^2=3+_6 3=0$ (*i.e.*, identity element) ;

$$\therefore \quad o(3)=2.$$

Similarly we can find orders of other elements. We have

$$o(4)=3, \; o(5)=6.$$

Example 3. *In the infinite multiplicative group of non-zero rational numbers, the order of every element except the elements* 1 *and* -1 *is infinite.*

We have $(-1)^1=-1$, $(-1)^2=1$ (identity element).

$\therefore \quad o(-1)=2.$

Now $2^1=2$, $2^2=4$, $2^3=8$ and so on. Thus there exists no positive integer n such that $2^n=1$ (identity element). Therefore $o(2)$ is infinite.

Example 4. *In the additive group of integers the order of every element except* 0 *is infinite.*

0 is the identity element. Therefore $o(0)=1$.

Now $1 \in \mathbf{I}$. We have $1(1)=1$, $2(1)=1+1=2$,

$3(1)=1+1+1=3$ and so on. Thus there exists no positive integer n such that $n(1)=0$ (identity element). Therefore $o(1)$ is infinite.

Note. In an infinite group elements may be of finite as well as of infinite order. We shall now prove some important results on the order of an element of a group.

Theorem 1. *The order of every element of a finite group is finite and is less than or equal to the order of the group.*

(Meerut 1971 ; Rajasthan 77)

Proof. Let G be a finite group, the composition being denoted multiplicatively. Let $a \in G$. Consider all positive integral powers of a *i.e.*, $a, a^2, a^3, a^4, \ldots$. All these are elements of G, by closure axiom. Since G has a finite number of elements, therefore all these integral powers of a cannot be distinct elements of G. Let us suppose that $a^r=a^s$ $(r > s)$.

Now $a^r=a^s \Rightarrow a^r a^{-s}=a^s a^{-s}$ $\quad [\because \; a^{-s} \in G]$

$$\Rightarrow a^{r-s}=a^0 \Rightarrow a^{r-s}=e$$

$$\Rightarrow a^m=e \text{ where } m=r-s.$$

Since $r>s$, therefore m is a positive integer. Thus there exists a positive integer m such that $a^m=e$.

Now we know that every set of positive integers has a least member. Therefore the set of all those positive integers m such that $a^m = e$ has least member, say n. Thus there exists a least positive integer n such that $a^n = e$. Therefore $o(a)$ is finite.

Now to prove that $o(a) \leqslant o(G)$.

Let $o(a) = n$ where $n > o(G)$. Since $a \in G$, therefore by closure property $a, a^2, \ldots, a^n$ are elements of G. No two of these are equal. For if possible, let $a^r = a^s$, $1 \leqslant s < r \leqslant n$. Then $a^{r-s} = e$. Since $0 < r-s < n$, therefore $a^{r-s} = e$ implies that the order of a is less than n. This is a contradiction. Hence $a, a^2, \ldots, a^n$ are n distinct elements of G. Since $n > o(G)$, therefore this is not possible. Hence we must have $o(a) \leqslant o(G)$.

Theorem 2. *The order of an element of a group is the same as that of its inverse* a^{-1}. **(Meerut 1968 ; Sagar 77)**

Proof. Let n and m be the orders of a and a^{-1} respectively.

We have $o(a) = n \Rightarrow a^n = e$ (identity element)

$$\Rightarrow (a^n)^{-1} = e^{-1} \Rightarrow (a^{-1})^n = e$$
$$\Rightarrow o(a^{-1}) \leqslant n \Rightarrow m \leqslant n.$$

Also $o(a^{-1}) = m \Rightarrow (a^{-1})^m = e$

$$\Rightarrow (a^m)^{-1} = e \Rightarrow a^m = e \quad [\because \; b^{-1} = e \Rightarrow b = e]$$
$$\Rightarrow o(a) \leqslant m \Rightarrow n \leqslant m.$$

Now $m \leqslant n$ and $n \leqslant m \Rightarrow m = n$.

If the order of a is infinite, then the order of a^{-1} cannot be finite. Because $o(a^{-1}) = m \Rightarrow o(a) \leqslant m \Rightarrow o(a)$ is finite. Therefore if the order of a is infinite, then the order of a^{-1} must also be infinite.

Theorem 3. *The order of any integral power of an element a cannot exceed the order of a.* **(Lucknow 1969 ; Garhwal 76)**

Proof. Let a^k be any integral power of a. Let $o(a) = n$.

Now $o(a) = n \Rightarrow a^n = e$ (identity element)

$$\Rightarrow (a^n)^k = e^k \Rightarrow a^{nk} = e$$
$$\Rightarrow (a^k)^n = e \Rightarrow o(a^k) \leqslant n.$$

Theorem 4. *If the element a of a group G is of order n, then* $a^m = e$ *iff n is divisor of m.* **(Lucknow 1980)**

Proof. Let n be a divisor of m. Then there exists an integer q such that $nq = m$.

Now $a^m = a^{nq} = (a^n)^q = e^q \quad [\because \; o(a) = n \Rightarrow a^n = e]$

$$= e.$$

Conversely let $a^m = e$.

Since m is an integer and n is a positive integer, therefore by division algorithm there exist integers q and r such that

$$m=nq+r \quad \text{where } 0 \leqslant r < n.$$

Now $a^m=a^{nq+r}=a^{nq}a^r=(a^n)^q a^r$

$$=e^q a^r \qquad [\because a^n=e]$$

$$=ea^r=a^r.$$

$\therefore$ $a^m=e \Rightarrow a^r=e.$

Since $0 \leqslant r < n$, therefore $a^r=e \Rightarrow r$ must be equal to zero because otherwise $o(a)$ will not be equal to n. If $o(a)=n$, then there will exist no positive integer $r < n$ such that $a^r=e$.

$\therefore$ $m=nq \Rightarrow n$ is a divisor of m.

Theorem 5. *The orders of the elements a and $x^{-1}ax$ are the same where a, x are any two elements of a group.* **(Gorakhpur 1970)**

Proof. Let n and m be the orders of a and $x^{-1}ax$ respectively.

Now $(x^{-1}ax)^2=(x^{-1}ax)(x^{-1}ax)=x^{-1}a(xx^{-1})ax$

$$=x^{-1}(ae)ax=x^{-1}aax=x^{-1}a^2x.$$

In general, we get

$$(x^{-1}ax)^n=x^{-1}a^n x$$

$$=x^{-1}ex \qquad [\because o(a)=n \Rightarrow a^n=e]$$

$$=x^{-1}x=e.$$

$\therefore$ $o(x^{-1}ax) \leqslant n \Rightarrow m \leqslant n.$

Again $o(x^{-1}ax)=m \Rightarrow (x^{-1}ax)^m=e$

$\Rightarrow x^{-1}a^m x=e \Rightarrow x^{-1}a^m x=x^{-1}x$

$\Rightarrow a^m x=x$ (by left cancellation law)

$\Rightarrow a^m x=ex$

$\Rightarrow a^m=e$ (by right cancellation law)

$\Rightarrow o(a) \leqslant m \Rightarrow n \leqslant m.$

Finally $m \leqslant n$, $n \leqslant m \Rightarrow m=n.$

Cor. *Order of ab is the same as that of ba where a and b are any elements of a group.* **(Banaras 1970)**

Proof. We have $a^{-1}(ab)a=(a^{-1}a)(ba)=e(ba)=ba.$

Thus $ba=a^{-1}(ab)a$

$\Rightarrow$ order of ba = order of $a^{-1}(ab)a$

$\Rightarrow$ order of ba = order of ab $\quad [\because o(x^{-1}ax)=o(a)]$

Theorem 6. *If a is an element of order n and p is prime to n, then a^p is also of order n.*

Proof. Let m be the order of a^p.

Now $o(a)=n \Rightarrow a^n=e \Rightarrow (a^n)^p=e^p=e$

$$\Rightarrow (a^p)^n=e \Rightarrow o(a^p) \leqslant n \Rightarrow m \leqslant n.$$

Since p, n are relative primes, there exist integers x and y such that $px+ny=1$. [Note]

$\therefore$ $a=a^1=a^{px+ny}=a^{px}\,a^{ny}=a^{px}\,(a^n)^y=a^{px}\,e^y=a^{px}\,e=a^{px}=(a^p)^x.$

Now $a^m=[(a^p)^x]^m=(a^p)^{mx}=[(a^p)^m]^x$

$=e^x$ [$\because$ $o(a^p)=m \Rightarrow (a^p)^m=e$]

$=e.$

$\therefore$ $o(a) \leqslant m \Rightarrow n \leqslant m.$

Finally $m \leqslant n$ and $n \leqslant m \Rightarrow m=n.$

Solved Examples

Ex. 1. *Prove that if* $a^2=a$, $a \in G$, *then* $a=e$.

(Kolhapur 1973)

Solution. We have $a^2=a \Rightarrow aa=a$

$\Rightarrow aa=ae$ [$\because$ $ae=a$]

$\Rightarrow a=e$

[by left cancellation law in G]

Ex. 2. *Given* $axa=b$ *in* G, *find* x.

Solution. We have $axa=b \Rightarrow a^{-1}(axa)=a^{-1}b$

$\Rightarrow (a^{-1}a)(xa)=a^{-1}b \Rightarrow e(xa)=a^{-1}b \Rightarrow xa=a^{-1}b$

$\Rightarrow (xa)a^{-1}=a^{-1}ba^{-1} \Rightarrow x(aa^{-1})=a^{-1}ba^{-1} \Rightarrow xe=a^{-1}ba^{-1}$

$\Rightarrow x=a^{-1}ba^{-1}.$

Ex. 3. *If* a *and* b *are any elements of a group* G, *then* $(bab^{-1})^n=ba^n b^{-1}$ *for any integer* n.

Solution. (i) $n=0$. We have $(ba\,b^{-1})^0=e$. [by definition]

Also $ba^0\,b^{-1}=be\,b^{-1}=bb^{-1}=e.$

$\therefore$ $(bab^{-1})^0=ba^0\,b^{-1}.$

(ii) $n>0$. We have $(bab^{-1})^1=ba\,b^{-1}=ba^1\,b^{-1}$, [$\because$ $a^1=a$]

Thus the result is true for $n=1$.

Let us suppose that the result is true for $n=k$ *i.e.*, suppose

$$(ba\,b^{-1})^k=ba^k\,b^{-1}.$$

Then $(ba\,b^{-1})^{k+1}=(ba\,b^{-1})^k\,(ba\,b^{-1})^1=ba^k\,b^{-1}\,ba\,b^{-1}$

$=ba^k\,eab^{-1}=ba^k\,ab^{-1}=ba^{k+1}\,b^{-1}.$

Therefore the result is true for $n=k+1$ if it was true for $n=k$. But we have seen that the result is true for $n=1$. Hence by mathematical induction it is true for all $n>0$.

(iii) $n<0$. Let $n=-m$ where $m>0$.

Then $(ba\,b^{-1})^n=(ba\,b^{-1})^{-m}=[(ba\,b^{-1})^m]^{-1}=(ba^m\,b^{-1})^{-1}$

$=(b^{-1})^{-1}\,(a^m)^{-1}\,b^{-1}=ba^{-m}\,b^{-1}=ba^n\,b^{-1}.$

Ex 4. *Prove that if G is an abelian group, then for all $a, b \in G$ and all integers n, $(ab)^n = a^n b^n$.*

Solution. (i) $n=0$. We have $(ab)^0=e$ [by definition]

Also $a^0 b^0=ee=e$.

$\therefore$ $(ab)^0=a^0 b^0$.

(ii) $n > 0$. If $n=1$, then $(ab)^1=ab=a^1b^1$.

Now suppose for $n=k$, $(ab)^k=a^k b^k$.

Then
$$(ab)^{k+1}=(ab)^k (ab)=a^k b^k ab$$
$$=a^k ab^k b \quad [\because G \text{ is abelian} \Rightarrow b^ka=ab^k]$$
$$=a^{k+1} b^{k+1}.$$

Thus the result is true for $n=k+1$ if it was true for $n=k$. But it is true for $n=1$. Hence by mathematical induction for all $n>0$, $(ab)^n=a^n b^n$.

(ii) $n<0$. Let $n=-m$ where m is a positive integer.

Then $(ab)^n=(ab)^{-m}=[(ab)^m]^{-1}=(a^m b^m)^{-1}$
$$=(b^m a^m)^{-1} \quad [\because G \text{ is abelian} \Rightarrow a^m b^m=b^m a^m]$$
$$=(a^m)^{-1} (b^m)^{-1} \quad [\because (ab)^{-1}=b^{-1} a^{-1}]$$
$$=a^{-m} b^{-m}=a^n b^n.$$

Note. Since $a, b \in G$, therefore $a^m, b^m \in G$.

Also G is abelian. Therefore $a^m b^m=b^m a^m$.

Ex 5. *Prove that if for every element a in a group G, $a^2=e$, then G is an abelian group.*

(G.N.D.U. Amritsar 1982; Nagarjuna 79; Madras 78; Gujrat 78)

Solution. Let a and b be any two elements of the group G. Then ab is also an element of G. Therefore $(ab)^2=e$.

Now $(ab)^2=e \Rightarrow (ab)(ab)=e \Rightarrow (ab)^{-1}=ab$
$$\Rightarrow b^{-1} a^{-1}=ab. \quad \ldots(1)$$

But $a^2=e \Rightarrow aa=e \Rightarrow a^{-1}=a$.

Similarly $b^2=e \Rightarrow b^{-1}=b$.

Therefore from (1), we get $ba=ab$. Thus we have $ab=ba$ $\forall$ $a, b \in G$. Therefore G is an abelian group.

Ex. 6. *Prove that a group G is abelian if every element of G except the identity element is of order two.*

(Punjab 1970; Meerut 79; Osmania 72)

Solution. Identity element e is of order 1. But $e^2=e$ Since every other element is of order two, therefore we have
$$a^2=e \;\forall\; a \in G.$$

Now proceed as in Ex. 5.

Ex. 7. *Show that if every element of a group G is its own inverse, then G is abelian.* **(Banaras 1970; Vikram 76; Raj. 78)**

Solution. Let a and b be any two elements of G. Then ab is also an element of G. Therefore $(ab)^{-1}=ab$ as it is given that every element is its own inverse.

Now $(ab)^{-1}=ab \Rightarrow b^{-1}a^{-1}=ab$

$\Rightarrow ba=ab.$ $[\because a^{-1}=a, b^{-1}=b]$

Thus we have $ab=ba \ \forall \ a, b \in G$ Therefore G is an abelian group.

Ex. 8. *Show that if a, b are any two elements of a group G, then* $(ab)^2=a^2b^2$ *if and only if G is abelian.*

(Meerut 1977, 81, 82; Nagpur 75; Kanpur 80; Madras 83; Sagar 77)

Solution. Suppose G is abelian.

Then $(ab)^2=(ab)\,(ab)=a\,(ba)\,b$

$=a\,(ab)\,b$ $[\because G$ is abelian $\Rightarrow ab=ba]$

$=(aa)\,(bb)=a^2b^2.$

Conversely, let a, b be any two elements of G.

Then $(ab)^2=a^2b^2 \Rightarrow (ab)\,(ab)=(aa)\,(bb) \Rightarrow a\,(ba)\,b=a\,(ab)\,b$

$\Rightarrow (ba)\,b=(ab)\,b$ [by left cancellation law]

$\Rightarrow ba=ab$ [by right cancellation law]

$\Rightarrow G$ is abelian.

Ex. 9. *Prove that a group G is abelian if*

$b^{-1}a^{-1}ba=e \ \forall \ a, b \in G.$

Solution. We have

$b^{-1}a^{-1}ba=e \Rightarrow (b^{-1}a^{-1})\,(ba)=e$

$\Rightarrow (b^{-1}a^{-1})^{-1}=ba$ $[\because ab=e \Rightarrow a^{-1}=b]$

$\Rightarrow (a^{-1})^{-1}\,(b^{-1})^{-1}=ba$ $[\because (ab)^{-1}=b^{-1}a^{-1}]$

$\Rightarrow ab=ba$ $[\because (a^{-1})^{-1}=a]$

$\Rightarrow G$ is abelian.

Ex. 10. *If G is a group of even order, prove that it has an element* $a\neq e$ *satisfying* $a^2=e$. **(Meerut 1983 P)**

Solution. Let G be a group of even order $2n$, where n is a positive integer. We shall prove that G must have an element $a\neq e$ such that $a^{-1}=a$. We shall prove it by contradiction.

Suppose G has no element, other than the identity element e, which is its own inverse. Now in a group every element possesses a unique inverse. The identity element e is its own inverse. Further if b is the inverse of c, then c is the inverse of b. So excluding the identity element e, the remaining $2n-1$ elements of G must be divided into pairs of two such that each pair consists of an element and its inverse. But we cannot do so because the odd integer $2n-1$ is not divisible by 2. Hence our initial assumption is wrong.

So in G there is an element $a \neq e$ such that

$$a = a^{-1} \Rightarrow aa = a^{-1}a \Rightarrow a^2 = e.$$

Ex. 11. (*a*) *What is the order of an* n*-cycle ?*

(*b*) *What is the order of the product of the disjoint cycles of lengths* $m_1, m_2, \ldots, m_k$ *?*

(*c*) *How do you find the order of a given permutation ?*

Solution. (a) Let $f = (1\ 2\ 3 \ldots n)$ be a cycle of length n.

Then f^2 moves every symbol two places along. Similarly f^3 will move every symbol three places along and f^n will move every symbol n places along. Thus $f^n = (1)\,(2)\,(3) \ldots (n)$ *i.e.*, identity permutation. Therefore order of f is n.

In particular if $f = (1\ 2\ 3\ 4\ 5)$, then $f^2 = (1\ 3\ 5\ 2\ 4)$, $f^3 = (1\ 4\ 2\ 5\ 3)$, $f^4 = (1\ 5\ 4\ 3\ 2)$, $f^5 = (1)\,(2)\,(3)\,(4)\,(5) =$ identity permutation.

$\therefore$ Order of f is 5.

(b) Suppose a permutation f is the product of disjoint cycles of lengths $m_1, m_2, \ldots, m_k$. As can be easily seen the order of f will be the L. C. M. (least common multiple) of the integers $m_1, m_2, \ldots, m_k$.

(c) To find the order of a permutation we should first express it as the product of disjoint cycles and then we should apply the rule given in part (b) of this question.

Ex. 12. *If a group G has four elements, show that it must be abelian.* **(Allahabad 1980; Kanpur 69; Delhi Hon's 69)**

Solution. Let $G = \{e, a, b, c\}$ be a group of order four. Here e is the identity element. The identity element e is its own inverse. There must be at least one more element in G which is its own inverse.

Let $a^{-1} = a$. If $b^{-1} = b$ and $c^{-1} = c$, then definitely G is abelian.

[See Ex. 7].

If $b^{-1} = c$, then $c^{-1} = b$ and we have

$$bc = e = cb. \quad \text{Also } a^{-1} = a \Rightarrow aa = e.$$

In this case the composition table for G will be as follows :

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	a	e
c	c	b	e	a

Note that ab would have been either equal to b or equal to c. Since $ab=b \Rightarrow a=e$, therefore ab must be equal to c. Then ac will be equal to b.

Now we can easily complete the table since in each column each element must be distinct and in each row each element must be distinct. From the table we see that composition in G is commutative. Therefore G is abelian.

Ex. 13. *If G is a group such that $(ab)^m=a^m b^m$ for three consecutive integers m for all $a, b \in G$, show that G is abelian.*

(I. A. S. 1970; Kurushetra 69; Meerut 69, 80)

Solution. Let a, b be any two elements of G. Suppose m, $m+1$, $m+2$ are three consecutive integers such that $(ab)^m=a^m b^m$, $(ab)^{m+1}=a^{m+1}\, b^{m+1}$ and $(ab)^{m+2}=a^{m+2}\, b^{m+2}$.

We have $(ab)^{m+2}=(ab)^{m+1}\,(ab)$

$\Rightarrow\ a^{m+2}\, b^{m+2}=a^{m+1}\, b^{m+1}\,(ab)$ (Given)

$\Rightarrow\ a^{m+1}\, ab^{m+1}\, b=a^{m+1}\, b^{m+1}\, ab$

$\Rightarrow\ ab^{m+1}=b^{m+1}\, a$ [by left and right cancellation laws]

$\Rightarrow\ a^m\,(ab^{m+1})=a^m\,(b^{m+1}\, a) \Rightarrow a^{m+1}\, b^{m+1}=a^m b^m\,(ba)$

$\Rightarrow\ (ab)^{m+1}=(ab)^m\, ba \Rightarrow (ab)^m\,(ab)=(ab)^m\, ba$

$\Rightarrow\ ab=ba$ [by left cancellation law]

$\Rightarrow\ G$ is abelian.

Ex. 14. *In a group, if $ba=a^m b^n$, prove that the elements $a^m b^{n-2}$, $a^{m-2}\, b^n$, ab^{-1} have the same order.* **(Raj. 1978)**

Solution. We have $a^m b^{n-2}=a^m b^n b^{-2}$

$=bab^{-2}$ $[\because\ ba=a^m b^n]$

$=ba\, b^{-1}\, b^{-1}=(b^{-1})^{-1}\,(ab^{-1})\, b^{-1}$.

Now we know that in a group $o\,(a)=o(x^{-1}\, ax)$, where a, x are any two elements of the group.

$\therefore\ \ o\,(a^m\, b^{n-2})=o\,[(b^{-1})^{-1}\,(ab^{-1})\, b^{-1}]=o\,(ab^{-1})$. ...(1)

Further $a^{m-2}\, b^n=a^{-2}\, a^m b^n=a^{-2}\, ba=a^{-2}\, ba^{2-1}=a^{-2}\, ba^{-1}a^2$

$$=(a^2)^{-1}(ba^{-1})\,a^2.$$

$$\therefore\quad o(a^{m-2}b^n)=o\,[(a^2)^{-1}(ba^{-1})\,a^2]=o\,(ba^{-1})$$

$$=o\,[(ba^{-1})^{-1}],\text{ since } o\,(a^{-1})=o\,(a)$$

$$=o\,[(a^{-1})^{-1}\,b^{-1}]=o\,(ab^{-1}). \qquad \ldots(2)$$

From (1) and (2), we get $o\,(a^m b^{n-2})=o\,(ab^{-1})=o\,(a^{m-2}b^n)$.

Ex. 15. *If in the group G, $a^5=e$, $aba^{-1}=b^2$ for $a, b \in G$ find $o\,(b)$.* **(Meerut 1976)**

Solution. We have $(ab\,a^{-1})^2=ab\,a^{-1}\,ab\,a^{-1}=ab^2\,a^{-1}$

$$=aa\,ba^{-1}\,a^{-1} \qquad [\because\ ab\,a^{-1}=b^2]$$

$$=a^2\,ba^{-2}.$$

$$\therefore\quad (ab\,a^{-1})^4=\{(ab\,a^{-1})^2\}^2=(a^2\,ba^{-2})^2=a^2\,ba^{-2}\,a^2\,ba^{-2}$$

$$=a^2\,b^2\,a^{-2}=a^2\,ab\,a^{-1}\,a^{-2}=a^3\,ba^{-3}.$$

$$\therefore\quad (a\,ba^{-1})^8=\{(ab\,a^{-1})^4\}^2=(a^3\,ba^{-3})^2=a^3ba^{-3}a^3ba^{-3}=a^3\,b^2\,a^{-3}$$

$$=a^3\,ab\,a^{-1}\,a^{-3}=a^4\,ba^{-4}.$$

$$\therefore\quad (a\,ba^{-1})^{16}=\{(ab\,a^{-1})^8\}^2=(a^4\,ba^{-4})^2=a^4ba^{-4}a^4ba^{-4}=a^4\,b^2a^{-4}$$

$$=a^4\,a\,ba^{-1}\,a^{-4}=a^5\,ba^{-5}$$

$$=eb\,e \qquad [\because\ a^5=e \text{ and so } a^{-5}=e]$$

$$=b.$$

Thus $(aba^{-1})^{16}=b$.

$$\therefore\quad (b^2)^{16}=b \qquad [\because\ ab\,a^{-1}=b^2]$$

$$\Rightarrow\quad b^{32}=b$$

$$\Rightarrow\quad b^{31}=e.$$

Since $b^m=e \Rightarrow o\,(b)\mid m$, therefore $o\,(b)\mid 31$.

But 31 is a prime integer. Therefore $o\,(b)=1$ or 31.

So if $b=e$, then $o\,(b)=1$ and if $b\neq e$, then $o\,(b)=31$.

Ex. 16. *If G is a finite abelian group with elements $a_1, a_2, \ldots, a_n$, prove that $a_1a_2\ldots a_n$ is an element whose square is the identity.*

Solution. We have $(a_1a_2\ldots a_n)^2=(a_1a_2\ldots a_n)\,(a_1a_2\ldots a_n)$. ...(1)

Now each element in a group has a unique inverse. Therefore each of $a_1, a_2, \ldots, a_n$ is the inverse of exactly one of them. So associating each of $a_1, a_2, \ldots, a_n$ with its inverse, the relation (1) becomes

$$(a_1a_2\ldots a_n)^2=(a_1a_1^{-1})\,(a_2a_2^{-1})\ldots(a_na_n^{-1})=eee\ldots\text{upto } n \text{ times}=e.$$

Ex. 17. *Show that the equation $x^2ax=a^{-1}$ is solvable for x in a group G if and only if a is the cube of some element in G.*

Solution. Suppose $x^2ax=a^{-1}$ is solvable in G. Then there exists an element $c \in G$ such that $c^2ac=a^{-1}$.

Now $c^2ac=a^{-1} \Rightarrow ccac=a^{-1} \Rightarrow c\,(ca)\,ca=a^{-1}\,a \Rightarrow c\,(ca)\,(ca)=e \Rightarrow (ca)\,(ca)=c^{-1} \Rightarrow (ca)\,(ca)\,c=c^{-1}\,c \Rightarrow (ca)\,(ca)\,c=e \Rightarrow (ca)\,(ca)\,ca=ea \Rightarrow (ca)^3=a \Rightarrow a$ is the cube of some element $ca \in G$.

Conversely, let $a=b^3$ for some $b \in G$. Then $x=b^{-2}$ is a solution of $x^2ax=a^{-1}$. For if $x=b^{-2}$ and $a=b^3$, then $x^2ax=b^{-4}b^3b^{-2}=b^{-3}=(b^3)^{-1}=a^{-1}$. Thus $x=b^{-2}$ is a solution of $x^2ax=a^{-1}$.

Ex. 18. *If in a group G, $xy^2=y^3x$ and $yx^2=x^3y$, then show that $x=y=e$ where e is the identity of G.* **(Nagarjuna 1980)**

Solution. We have $xy^2=y^3x \Rightarrow x^2y^2=xy^3x$

$\Rightarrow x^2y=xy^3xy^{-1}=xy^2yxy^{-1}=y^3xyxy^{-1}$ ...(1)

Again $yx^2=x^3y \Rightarrow yx^2=xx^2y$

$\Rightarrow yx^2=xy^3xyxy^{-1}$, substituting for x^2y from (1)

$\Rightarrow x^2=y^{-1}\,xy^3xyxy^{-1}$

$\Rightarrow x^2y=y^{-1}xy^3xyx.$ ...(2)

From (1) and (2), we get

$$y^3xyxy^{-1}=y^{-1}xy^3xyx$$

$\Rightarrow y^4xyx=xy^3xyxy=xy^2yxyxy=y^3xyxyxy.$

Cancelling y^3 from both sides, we get

$$yxyx=xyxyxy$$

$\Rightarrow (yx)^2=(xy)^3.$ ...(3)

Since the given relations are symmetrical in x and y, therefore interchanging x and y in (3), we get

$$(xy)^2=(yx)^3. \qquad ...(4)$$

Now from (3) and (4), we have

$$(xy)^2=(yx)^3=(yx)^2\,(yx)=(xy)^3\,(yx).$$

Cancelling $(xy)^2$ from both sides. we get

$$e=(xy)\,(yx)=xy^2x$$
$$\Rightarrow x^{-2}=y^2.$$

Now $xy^2=y^3x \Rightarrow xx^{-2}=yx^{-2}x$

$$\Rightarrow x^{-1}=yx^{-1} \Rightarrow y=e.$$

Again $yx^2=x^3y \Rightarrow ex^2=x^3e \Rightarrow x^2=x^3 \Rightarrow x=e.$

Ex. 19. *Let G be a group and let $a \in G$ be of finite order n. Then for any integer k, we have $o\,(a^k)=\dfrac{n}{(n,\,k)}$ where (n, k) denotes the H.C.F. of n and k.*

Solution. Let $(n, k)=m$. Then we have $n=pm$, $k=qm$ for some integers p and q such that $(p, q)=1$. Let $o(a^k)=l$.

We have $o(a^k)=l \Rightarrow (a^k)^l=e \Rightarrow a^{kl}=e$

$\Rightarrow n \mid kl$ $[\because\ o(a)=n;\ \therefore\ a^{kl}=e \Rightarrow n \mid kl]$

$\Rightarrow pm \mid qml \Rightarrow p \mid ql$

$\Rightarrow p \mid l$ [$\because$ p and q are relatively prime]

Again $(a^k)^p=(a^{qm})^p=a^{qmp}=a^{qn}=(a^n)^q=e^q=e$.

Therefore $o(a^k) \mid p$ *i.e.*, $l \mid p$.

Now $l \mid p$ and $p \mid l \Rightarrow l=p$.

$$\therefore\ o(a^k)=p=\frac{n}{m}=\frac{n}{(n, k)}.$$

Ex. 20. General law of commutativity of the elements of a group. *If in a group* G, $a_1, a_2, \ldots, a_n$ *is any system of* n *elements,* **commutative in pairs,** *and*

$$\begin{pmatrix} 1 & 2 \ldots n \\ i_1 & i_2 \ldots i_n \end{pmatrix}$$

is any permutation of the set of n *objects* $1, 2, \ldots, n$, *then*

$$a_1 a_2 \ldots a_n = a_{i_1} a_{i_2} \ldots a_{i_n}.$$

(Punjab 1970 ; Meerut 70)

Solution. We shall prove the result by induction method. Suppose the result is true for products of $n-1$ or less elements. Then we shall show that it is also true for products of n elements. Two cases arise :

Case I. Let $i_n=n$. In this case we have

$a_1 a_2 \ldots a_n=(a_1 a_2 \ldots a_{n-1})\, a_n$ [$\because$ composition in G is associative]

$=(a_{i_1} a_{i_2} \ldots a_{i_{n-1}})\, a_n$ [$\because$ according to supposition the result is true for $n-1$ elements]

$=(a_{i_1} a_{i_2} \ldots a_{i_{n-1}})\, a_{i_n}$ $[\because\ n=i_n]$

$=a_{i_1} a_{i_2} \ldots a_{i_{n-1}} a_{i_n}$ [by associativity]

Case II. Let $i_n=k$ where $k < n$. Then

$a_1 a_2 \ldots a_{k-1}\, a_k\, a_{k+1} \ldots a_n$

$=(a_1 a_2 \ldots a_{k-1})\,(a_k\, a_{k+1} \ldots a_n)$ [by associativity]

$=(a_1 a_2 \ldots a_{k-1})\,(a_{k+1} \ldots a_n a_k)$ [by supposition]

$=(a_1 a_2 \ldots a_{k-1})\,[(a_{k+1} \ldots a_n)\, a_k]=[(a_1 a_2 \ldots a_{k-1})\,(a_{k+1} \ldots a_n)]\, a_k$

$=(a_1 a_2 \ldots a_{k-1}\, a_{k+1} \ldots a_n)\, a_k$

$=(a_{i_1} a_{i_2} \ldots a_{i_{n-1}})\, a_{i_n}=a_{i_1} a_{i_2} \ldots a_{i_{n-1}} a_{i_n}.$

Now obviously the result is true for products of 2 elements. Hence the proof is complete by induction.

Exercises

1. Define the order of an element in (i) an additive group, (ii) a multiplicative group. What is the order of the residue class [3] in the multiplicative group of non-zero residue classes modulo 5 ? **Ans. 4.** **(Gorakhpur 1970)**

2. Distinguish between the order of a group and the order of an element in a group. Prove that if $a, x \in G$ then a and xax^{-1} have the same order in G. **(Gorakhpur 1970; Meerut 70)**

3. Show that in a group G, we have

$$ab=e \Rightarrow a=b^{-1} \text{ and } b=a^{-1}.$$

4. Show that in a group G, we have

$$ab=a \text{ or } ba=a \Rightarrow b=e,$$

where e is the identity element of G. **(Kanpur 1969)**

5. If in a group G, the elements a and b commute, then prove that (i) a^{-1} and b^{-1} also commute. (ii) a^{-1} and b also commute, (iii) a and b^{-1} also commute.

6. In a group G, prove that $a^n=e$ for any integer n.

7. Find the solution of the equation $abxax=cbx$ in a group G, where a, b and c are given elements of G.

Ans. $x=b^{-1}a^{-1}cba^{-1}$.

8. Prove that, if a group has an even number of elements, then at least one element, apart from the identity element must equal its inverse.

9. Prove that in any group, e is the only element of order 1

10. Find the order of each element of the group

$$(\{0, 1, 2, 3, 4\}, +_5).$$ **(Gorakhpur 1970)**

Ans. $o(0)=1$ and each other element is of order 5.

11. Find the order of the permutation $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}$.

Ans. 3.

12. Let G be a group and let $a \neq e \in G$ be such that $a^p=e$, where p is a prime number. Prove that $o(a)=p$.

13. Show by means of an example that it is possible for the quadratic equation $x^2=e$ to have more than two solutions in some group G with identity e. **(Marathwada 1972; Meerut 81)**

14. If the elements a, b and ab of a group are each of order 2, prove that $ab=ba$. **(Meerut 1973)**

15. Let $a, b \in G$ be non-identity elements with $o(a)=5$ and $a^{-1}ba=b^2$. Find $o(b)$.

Ans. 31.

16. Give an example of an infinite group each element of which has a finite order. **(Poona 1970)**

[**Hint.** Let $G=\{z : z \in \mathbf{C}$ and $z^n=1$ for some positive integer $n\}$ *i.e.*, let G be the union of all n nth roots of unity where $n \in \mathbf{N}$. Then for multiplication of complex numbers G is the desired group].

17. Let $x \in G$, $x \neq e$. Show that $x \neq x^{-1}$ if and only if the order of x is greater than two, where G is any group.

(Poona 1973)

18. If a is an element of a group, prove that the integral powers of a form a multiplicative group. **(Meerut 1975)**

19. In S_3 give an example of two elements x, y such that $(xy)^2 \neq x^2y^2$. Here S_3 is the symmetric group of permutations of degree 3.

20. In S_3 show that there are four elements satisfying $x^2=e$ and three elements satisfying $y^3=c$.

21. In a group G let a be an element of finite order n. If a positive integer k is a divisor of n, prove that $o(a^k)=n/k$.

22. Show that the equation $xax=b$ is solvable for x in a group G if and only if ab is the square of some element in G.

23. Show that in an abelian group the product of two elements of finite order is again an element of finite order.

24. Prove that a group G is abelian if and only if

$$(ab)^{-1}=a^{-1}b^{-1} \ \forall \ a, b \in G.$$ **(Rohilkhand 1981)**

§ 18. Isomorphism of groups.

Now we shall discuss the very important concept of isomorphism of groups.

Isomorphic mapping. Definition.

(Agra 1977 ; Osmania 72 ; Rohilkhand 80 ; Poona 73)

Suppose G and G' are two groups, the composition in each being denoted multiplicatively. A mapping f of G into G' is said to be an isomorphic mapping of G into G' if

(*i*) *f is one-to-one i.e., distinct elements in G have distinct f-images in G',* (*ii*) *$f(ab)=f(a)\,f(b)\ \forall\ a, b \in G$ i.e., the image of the product is the product of the images.*

It should be noted that when we say that f is a mapping of G into G', we usually include in it the possibility that the mapping f may be onto G'. If an isomorphic mapping f of G into G' is onto G', then it is called an isomorphic mapping of G onto G'.

If $f : G \to G'$ be such that $f(ab)=f(a)\,f(b)\ \forall\ a, b \in G$, then we say that the mapping f **preserves the compositions** in G and G'. Here $f(ab)$ is an element of G' and it is the f-image of ab which is an element of G obtained on multiplying a and b. Further $f(a)$ and $f(b)$ are elements of G'. Also $f(a)\,f(b)$ is an element of G' obtained on multiplying $f(a)$ and $f(b)$.

Thus we can say that, a mapping f is said to be an isomorphic mapping of G into G' if it is one-to-one and if it preserves the compositions in G and G'.

If f is an isomorphic mapping of a group G into a group G', then f is also called an **isomorphism** of G into G'. If f is an isomorphism of G **onto** G', the group G' is called an **isomorphic image** of the group G. Also then we say that the group G is isomorphic to the group G'. Thus we can give the complete definition of isomorphic groups like this :

Isomorphic groups. Definition.

(Meerut 1978, 81, 82 P; Osmania 72; Agra 71; I.A.S. 70)

Suppose G and G' are two groups. Further suppose that the compositions in both G and G' have been denoted multiplicatively. Then we say that the group G is isomorphic to the group G' if there exists a one-to-one mapping f of G onto G' such that

$$f(ab)=f(a)\,f(b)\ \forall\ a, b \in G$$

i.e., the mapping f preserves the compositions in G and G'.

If the group G is isomorphic to the group G', symbolically we write $G \cong G'$. Another notation for isomorphism is $\approx$.

Note 1. In our definition of isomorphism of two groups G and G' we have denoted the compositions in G and G' both multiplicatively. The students can use different symbols to denote the compositions in G and G'. But there should be no confusion.

Note 2. If G is isomorphic to G', there may exist more than one isomorphisms of G onto G'. There may be many one-one onto functions from G to G'. But if there exists at least one function f which is one-one, onto and also preserves compositions, then G will be isomorphic to G'.

Note 3. If the group G is finite, then G can be isomorphic to G' only if G' is also finite and the number of elements in G is equal to the number of elements in G'. Otherwise there will exist no mapping f from G to G' which is one-one as well as onto.

Note 4. If the group G is isomorphic to the group G', then we say that the groups G and G' are **abstractly identical.** From the point of view of abstract algebra we shall regard them as one group and not as two different groups.

Example 1. *If* **R** *is the additive group of real numbers and* $\mathbf{R}_+$ *the multiplicative group of positive real numbers, prove that the mapping* $f : \mathbf{R} \to \mathbf{R}_+$ *defined by* $f(x)=e^x \ \forall \ x \in \mathbf{R}$ *is an isomorphism of* **R** *onto* $\mathbf{R}_+$. **(Gujrat 1970, Meerut 81, 82 P; Delhi 70, Kolhapur 73, Gorakhpur 70, Madurai 78).**

Solution. If x is any real number positive, zero or negative, then e^x is always a positive, real number. Also e^x is unique. Therefore if $f(x)=e^x$, then $f : \mathbf{R} \to \mathbf{R}_+$.

f is one-to-one.

Let $x_1, x_2 \in \mathbf{R}$. Then $f(x_1)=f(x_2)$

$\Rightarrow e^{x_1}=e^{x_2}$ [$\because$ by the def. of $f, f(x)=e^x$]

$\Rightarrow \log e^{x_1}=\log e^{x_2} \Rightarrow x_1 \log e=x_2 \log e \Rightarrow x_1=x_2$.

Thus two elements in **R** have the same f-image in $\mathbf{R}_+$ only if they are equal. Consequently distinct elements in **R** have distinct f-images in $\mathbf{R}_+$. Therefore f is one-to-one.

f is onto. Suppose y is any element of $\mathbf{R}_+$ *i e.*, y is any positive real number. Then $\log y$ is a real number *i.e.*, $\log y \in \mathbf{R}$.

Now $f(\log y)=e^{\log y}=y$. Thus $y \in \mathbf{R}_+ \Rightarrow$ that $\exists \log y \in \mathbf{R}$ such that $f(\log y)=y$. Therefore each element of $\mathbf{R}_+$ is the f-image of some element of **R**. Thus f is onto.

f preserves compositions in R and $\mathbf{R}_+$. Suppose x_1 and x_2 are any two elements of **R**. Then

$$f(x_1+x_2)=e^{x_1+x_2} \quad \text{[by definition of } f]$$

$$=e^{x_1}.e^{x_2}$$

$$=f(x_1).f(x_2) \quad [\because f(x_1)=e^{x_1} \text{ and } f(x_2)=e^{x_2}].$$

Thus f preserves compositions in **R** and $\mathbf{R}_+$. Here the composition in **R** is addition and the composition in $\mathbf{R}_+$ is multiplication. Therefore f is an isomorphism of **R** onto $\mathbf{R}_+$. Hence $\mathbf{R} \cong \mathbf{R}_+$.

Example 2. *Let* $\mathbf{R}_+$ *be the multiplicative group of all positive real numbers and* $\mathbf{R}$ *be the additive group of all real numbers. Show that the mapping* $g : \mathbf{R}_+ \to \mathbf{R}$ *defined by*

$g(x) = \log x \ \forall \ x \in \mathbf{R}_+$ *is an isomorphism.* **(Meerut 1976)**

Solution. If x is any positive real number, then $\log x$ is definitely a real number. Also $\log x$ is unique. Therefore if $g(x) = \log x$, then $g : \mathbf{R}_+ \to \mathbf{R}$.

g is one-to-one. Let $x_1, x_2 \in \mathbf{R}_+$ Then $g(x_1) = g(x_2)$

$$\Rightarrow \log x_1 = \log x_2 \Rightarrow e^{\log x_1} = e^{\log x_2} \Rightarrow x_1 = x_2.$$

Therefore g is one-to-one.

g is onto. Suppose y is any element of $\mathbf{R}$ *i.e.*, y is any real number. Then e^y is definitely a positive real number *i.e.*, $e^y \in \mathbf{R}_+$.

Now $g(c^y) = \log e^y = y$. Thus $y \in \mathbf{R} \Rightarrow$ that there exists $e^y \in \mathbf{R}_+$ such that $g(e^y) = y$. Therefore each element of $\mathbf{R}$ is the g-image of some element of $\mathbf{R}_+$. Thus g is onto.

g preserves compositions in $\mathbf{R}_+$ and $\mathbf{R}$. Suppose x_1 and x_2 are any two elements of $\mathbf{R}_+$. Then

$$\begin{aligned} g(x_1 x_2) &= \log(x_1 x_2) && [\text{by def. of } g] \\ &= \log x_1 + \log x_2 \\ &= g(x_1) + g(x_2) && [\text{by def. of } g] \end{aligned}$$

Thus g preserves compositions in $\mathbf{R}_+$ and $\mathbf{R}$. Here the composition in $\mathbf{R}_+$ is multiplication and the composition in $\mathbf{R}$ is addition. Therefore g is an isomorphism of $\mathbf{R}_+$ onto $\mathbf{R}$. Hence $\mathbf{R}_+ \cong \mathbf{R}$.

Ex. 3. *Show that the additive group of integers*

$$G = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$$

is isomorphic to the additive group

$$G' = \{\ldots, -3m, -2m, -1m, 0, 1m, 2m, 3m, \ldots\}$$

where m is any fixed integer not equal to zero. **(Agra 1971)**

Solution. If $x \in G$, then obviously $mx \in G'$. Let $f : G \to G'$ be defined by $f(x) = mx \ \forall \ x \in G$.

f is one-to-one. Let $x_1, x_2 \in G$. Then $f(x_1) = f(x_2)$

$$\begin{aligned} &\Rightarrow mx_1 = mx_2 && [\text{by def. of } f] \\ &\Rightarrow x_1 = x_2 && [\because m \neq 0] \end{aligned}$$

Therefore f is one-to-one.

f is onto. Suppose y is any element of G'. Then obviously $y/m \in G$. Also $f(y/m) = m(y/m) = y$. Thus $y \in G' \Rightarrow$ that there exists $y/m \in G$ such that $f(y/m) = y$. Therefore each element of G' is the f-image of some element of G. Hence f is onto.

Again if x_1 and x_2 are any two elements of G, then

$f(x_1+x_2)=m(x_1+x_2)$ [by def. of f]

$=mx_1+mx_2$ [by distributive law for integers]

$=f(x_1)+f(x_2)$. [by definition of f]

Thus f preserves compositions in G and G'. Therefore f is an isomorphic mapping of G onto G'. Hence G is isomorphic to G'.

Ex. 4. *Show that the set* **C** *of all complex numbers under addition is a group which is isomorphic to itself under the identity mapping as well as under the mapping which takes every complex number into its conjugate complex.*

Solution. The identity mapping f defined by $f: \mathbf{C} \to \mathbf{C}$ such that $f(z)=z \ \forall \ z \in \mathbf{C}$ is obviously one-one onto.

Also $f(z_1+z_2)=z_1+z_2=f(z_1)+f(z_2) \ \forall \ z_1, z_2 \in \mathbf{C}$.

$\therefore$ the identity mapping f is an isomorphism of **C** onto **C**.

If $z=x+iy$ is any complex number, then $\bar{z}=x-iy$ is called the conjugate complex of z.

Let $g: \mathbf{C} \to \mathbf{C}$ be such that $g(z)=\bar{z} \ \forall \ z \in \mathbf{C}$.

Let $z_1, z_2 \in \mathbf{C}$. Then $g(z_1)=g(z_2) \Rightarrow \bar{z}_1=\bar{z}_2$

$\Rightarrow \overline{(\bar{z}_1)}=\overline{(\bar{z}_2)} \Rightarrow z_1=z_2$.

Therefore g is one-to-one.

If $x+iy$ is any element of **C**, then $x-iy$ is also an element of **C**. Also $g[(x-iy)]=x+iy$. Therefore g is onto.

Further if $z_1, z_2 \in \mathbf{C}$, then $g(z_1+z_2)=\overline{(z_1+z_2)}=\bar{z}_1+\bar{z}_2$

$=g(z_1)+g(z_2)$.

Hence g is also an isomorphism of **C** onto **C**.

§ 19 Some Important Properties of Isomorphic mappings.

Let f be an isomorphic mapping of a group G into a group G'. Then we have the following important properties.

(i) *The f-image of the identity e of G is the identity of G' i.e., $f(e)$ is the identity of G'.* **(Kanpur 1970; Banaras 66; Allahabad 70; Raj. 78; Berhampur 77; Meerut 68; Rohilkhand 80; Lucknow 66)**

Proof. Let e be the identity of G and e' be the identity of G'. Let a be any element of G. Then $f(a) \in G'$.

Now $e' f(a)=f(a)$ [$\because$ e' is the identity of G']

$=f(ea)$ [$\because$ e is the identity of G]

$=f(e) f(a)$ [$\because$ f is an isomorphic mapping]

Now in the group G', we have

$e' f(a)=f(e) f(a)$

$\Rightarrow e'=f(e)$ [by right cancellation law in G']

$\therefore$ $f(e)$ is the identity of G'.

(ii) The f-image of the inverse of an element a of G is the inverse of the f-image of a i.e., $f(a^{-1})=[f(a)]^{-1}$.

(Kanpur 1970; Garhwal 76; Berhampur 77; Rohilkhand 80; Meerut 76)

Proof. Suppose e is the identity of G and e' is the identity of G'. Then $f(e)=e'$. Now let a be any element of G. Then $a^{-1}\in G$ and $aa^{-1}=e$. We have

$e'=f(e)=f(aa^{-1})=f(a)f(a^{-1})$ [$\because$ f is composition preserving]

Therefore $f(a^{-1})$ is the inverse of $f(a)$ in the group G'. Thus $f(a^{-1})=[f(a)]^{-1}$.

(iii) The order of an element a of G is equal to the order of its image f (a) **(Poona 73)**

Proof. Suppose e is the identity of G. Then $f(e)$ is the identity of G'. Let the order of a be finite and let it be equal to n.

Then $a^n=e \Rightarrow f(a^n)=f(e)$

$\Rightarrow f(aaa\ldots n \text{ times})=f(e)$

$\Rightarrow f(a)f(a)\ldots n \text{ times}=f(e)$

$\Rightarrow [f(a)]^n=f(e) \Rightarrow$ order of $f(a) \leqslant n$.

If now the order of $f(a)$ is m, then

$[f(a)]^m=f(e)$

$\Rightarrow f(a)f(a)f(a)\ldots m \text{ times}=f(e)$

$\Rightarrow f(aaa\ldots m \text{ times})=f(e)\Rightarrow f(a^m)=f(e)$

$\Rightarrow a^m=e$ [$\because$ f is one-one]

$\Rightarrow$ order of $a\leqslant m$.

Thus $m\leqslant n$ and $n\leqslant m \Rightarrow m=n$.

If the order of a is infinite, then the order of $f(a)$ cannot be finite. Because if the order of $f(a)$ is finite and is equal to m, then we have $a^m=e$. Therefore the order of a is finite. Thus we get a contradiction.

Hence in all cases order of a=order of $f(a)$.

Important Note Suppose we are to prove that a group G is isomorphic to another group G'. Then we should try to find a one-one mapping from G onto G' which also preserves compositions in G and G'. While forming such a mapping we should keep in mind the above three facts that an isomorphic mapping must preserve identities, inverses and orders.

Solved Examples

Ex. 1. *Show that the multiplicative group* $G=\{1, -1, i, -i\}$ *is isomorphic to the permutation group*

$G'=\{I, (abcd), (ac)(bd), (adcb)\}$
on four symbols a, b, c, d. **(Banaras 70)**

Solution. We note that 1 is the identity of G and I is the identity of G'. Therefore if f is to be an isomorphism of G onto G', then we must have $f(1)=I$ *i.e.*, identity must go to identity.

In the group G the orders of -1, i and $-i$ are 2, 4 and 4 respectively. While in the group G' the orders of $(abcd)$, $(ac)(bd)$, $(adcb)$ are 4, 2, and 4 respectively. In an isomorphic mapping only elements of equal order can be mapped on each other. Therefore we must have $f(-1)=(ac)(bd)$. Now the f-image of i can be taken either $(abcd)$ or $(adcb)$. Let us take $f(i)=(abcd)$ and $f(-i)=(adcb)$. It should be noted that the inverse of i in G is $-i$ while the inverse of $(abcd)$ in G' is $(adcb)$.

The mapping $f: G\to G'$ as defined above is obviously one-one onto. Now to show that f preserves compositions in G and G' we proceed as follows :

Let us write $A_1=1$, $A_2=-1$, $A_3=i$ and $A_4=-i$. We shall denote the corresponding f-images by writing B in place of A. So let us write $B_1=f(A_1)=I$, $B_2=f(A_2)=(ac)(bd)$, $B_3=f(A_3)=(abcd)$ and $B_4=f(A_4)=(adcb)$.

Now we shall form the composition tables for the two groups:

	A_1	A_2	A_3	A_4
A_1	A_1	A_2	A_3	A_4
A_2	A_2	A_1	A_4	A_3
A_3	A_3	A_4	A_2	A_1
A_4	A_4	A_3	A_1	A_2

	B_1	B_2	B_3	B_4
B_1	B_1	B_2	B_3	B_4
B_2	B_2	B_1	B_4	B_3
B_3	B_3	B_4	B_2	B_1
B_4	B_4	B_3	B_1	B_2

We see that the composition tables for G and G' are identical *i.e.*, if we replace A_1, A_2, A_3, A_4 by B_1, B_2, B_3, B_4 in the composition table for G, we reproduce the complete composition table for G'. Therefore the mapping f preserves compositions in G and G'. Obviously $f(A_iA_j)=B_iB_j=f(A_i)\,f(A_j)\ \forall\ A_i, A_j\in G$.

For example, $f(A_3A_4)=f(A_1)=B_1=B_3B_4=f(A_3)\,f(A_4)$.

$\therefore$ f is an isomorphism of G onto G'. Hence G is isomorphic to G'.

Note. The mapping $\phi : G \to G'$ defined by $\phi(1)=I$, $\phi(-1)=(ac)(bd)$, $\phi(i)=(adcb)$, $\phi(-i)=(abcd)$ is also an isomorphism of G onto G' as can be easily seen. Also it is obvious that in this case we can have only two isomorphisms of G onto G'.

Ex. 2. *Show that the multiplicative group* $G=\{1, \omega, \omega^2\}$ *is isomorphic to the permutation group* $G'=\{I, (abc), (acb)\}$ *on three symbols* a, b, c.

Solution. Proceed as in Ex. 1. Take the mapping $f : G \to G'$ defined as $f(1)=I$, $f(\omega)=(abc)$, $f(\omega^2)=(acb)$.

Ex. 3. *Show that the group* $G=(\{0, 1, 2, 3\}, +_4)$ *is isomorphic to the group* $G'=(\{1, 2, 3, 4\}, \times_5)$.

Solution. Here 0 is the identity of G and 1 is the identity of G'. Therefore if $f : G \to G'$ is to be an isomorphism of G onto G', we must have $f(0)=1$.

In the group G the orders of 1, 2, 3 are 4, 2 and 4 respectively. In the group G' the orders of 2, 3, 4 are 4, 4 and 2 respectively. Therefore we must take $f(2)=4$. Further let us take $f(1)=2$, $f(3)=3$.

The mapping f is obviously one-one onto. To show that f preserves compositions in G and G', we should proceed as in Ex. 1.

§ 20. The relation of Isomorphism in the set of all groups.

Theorem. *The relation of isomorphism in the set of all groups is an equivalence relation.* **(Punjab 1970; Gorakhpur 70; Lucknow 68 ; Banaras 69; Meerut 81, 84)**

Proof. We shall prove that the relation of isomorphism denoted by $\cong$ in the set of all groups is reflexive, symmetric and transitive.

Reflexive. If G is any group, then $G \cong G$. Let f be the identity mapping on G *i.e.*, let $f : G \to G$ such that $f(x)=x \ \forall \ x \in G$. Obviously f is one-one onto. Also if x, y are any elements of G, then $f(x)=x$ and $f(y)=y$.

$$\text{Also} \quad f(xy)=xy \qquad [\because f \text{ is identity mapping}]$$
$$=f(x)\, f(y).$$

$\therefore$ f is composition preserving also. Thus f is an isomorphism of G onto G.

Hence $\qquad G \cong G.$

Symmetric *i.e.*, $G \cong G' \Rightarrow G' \cong G$. Suppose a group G is isomorphic to another group G'. Let f be an isomorphism of G onto G'. Then f is one-one onto and preserves compositions in G

and G'. Since f is one-one onto, therefore it is inversible *i.e.*, f^{-1} exists. Also we know that the mapping f^{-1} is also one-one onto.

Now we shall show that $f^{-1} : G' \to G$ is also composition preserving. Let x', y' be any elements of G'. Then there exist elements $x, y \in G$ such that

$$f^{-1}(x') = x, \; f^{-1}(y') = y \qquad \ldots(1)$$

and $$f(x) = x', \; f(y) = y'. \qquad \ldots(2)$$

Now $$f^{-1}(x'y') = f^{-1}[f(x) f(y)] \qquad \text{[From (2)]}$$
$$= f^{-1}[f(xy)], \text{ since } f(xy) = f(x) f(y)$$
$$= xy \qquad \text{[by def. of } f^{-1}]$$
$$= f^{-1}(x') f^{-1}(y'). \qquad \text{[From (1)]}$$

$\therefore$ f^{-1} preserves compositions in G' and G.

Hence $$G' \cong G.$$

Transitive *i.e.*, $G \cong G', G' \cong G'' \Rightarrow G \cong G''$. Suppose G is isomorphic to G' and G' is isomorphic to G''. Further suppose that $f : G \to G'$ and $g : G' \to G''$ are the respective isomorphic mappings. We know that the composite mapping $g \circ f : G \to G''$ defined by

$$(g \circ f)(x) = g[f(x)] \; \forall \; x \in G$$

is also one-one onto if both f and g are one-one onto.

Further if x, y are any elements of G, then

$$(g \circ f)(xy) = g[f(xy)] \qquad \text{[by definition of } g \circ f]$$
$$= g[f(x) f(y)] \qquad [\because f \text{ is composition preserving}]$$
$$= g[f(x)] g[f(y)] \qquad [\because g \text{ is also an isomorphism}]$$
$$= [(g \circ f)(x)] [(g \circ f)(y)].$$

Hence $g \circ f$ preserves compositions in G and G''.

$\therefore$ $g \circ f$ is an isomorphism of G onto G'' and $G \cong G''$

Hence the relation of isomorphism in the set of all groups is an equivalence relation.

Note. The relation of isomorphism in the set of all groups will partition the set of all groups into disjoint equivalence classes. If G_1 is any group, then all the groups isomorphic to G_1 will form one equivalence class. If G_2 is another group not isomorphic to G_1, then all the groups isomorphic to G_2 will form another equivalence class, and so on.

§ 21. Transference of group structures.

Theorem. *If G is a group and G' is a set equipped with a composition (supposed denoted multiplicatively) and if there exists a one-one mapping f of G onto G' such that*

$$f(ab) = f(a) f(b) \; \forall \; a, b \in G,$$

then G' is also a group isomorphic to G for the composition in question.

Proof. To show that G' is a group we are to show that the composition in G' satisfies the group postulates.

Associativity. Suppose a', b', c' are any three elements of G'. Since f is an onto function, there exist three elements a, b, c of G such that $f(a)=a', f(b)=b', f(c)=c'$.

Now $(a'b')\,c'=[f(a)\,f(b)]\,f(c)$

$=[f(ab)]\,f(c)$ [$\because f$ preserves compositions]

$=f\,[(ab)\,c]$ [$\because f$ is composition preserving]

$=f\,[a\,(bc)]$ [by associativity in the group G]

$=f(a)\,[f(bc)]=f(a)\,[f(b)\,f(c)]=a'\,(b'\,c')$.

Therefore the composition in G' is associative.

Existence of Identity. Suppose e is the identity of G. Then $f(e)$ is an element of G'. If a' is any element of G', there exists an element $a \in G$ such that $f(a)=a'$.

Now $[f(e)]\,a'=[f(e)]\,[f(a)]=f(ea)=f(a)=a'$.

Also $a'\,[f(e)]=[f(a)]\,[f(e)]=f(ae)=f(a)=a'$.

$\therefore$ $f(e)$ which is the f-image of the identity e in G is the identity in G'.

Existence of Inverse. If a' is any element of G', there exists an element $a \in G$ such that $f(a)=a'$.

Now $a \in G \Rightarrow a^{-1} \in G$ where $a^{-1}a=e=aa^{-1}$.

Also $f(a^{-1})$ is an element of G'.

Now $[f(a^{-1})]\,a'=[f(a^{-1})]\,[f(a)]=f(a^{-1}a)=f(e)$.

Also $a'\,[f(a^{-1})]=[f(a)]\,[f(a^{-1})]=f(aa^{-1})=f(e)$.

$\therefore$ $f(a^{-1})$ is the inverse of a' in G'. Thus every element of G' possesses inverse.

Exercises

1. Show that the mapping

$$f: G \rightarrow G' \text{ defined by } f(x)=2x \ \forall\ x \in G$$

is an isomorphism of G onto G' where G is the additive group of integers and G' is the additive group of even integers including zero.

2. Show that the additive group G of all integers is isomorphic to the multiplicative group

$$G'=\{\ldots, 3^{-3}, 3^{-2}, 3^{-1}, 3^{0}, 3^{1}, 3^{2}, 3^{3}, \ldots\}.$$

3. Show that the multiplicative group $G=\{1, -1\}$ is isomorphic to the permutation group $G'=\{I, (ab)\}$.

4. (*i*) What do you understand by the statement :

"When two groups are isomorphic, then in some sense they are equal". **(Allahabad 80)**

(*ii*) Are any two finite groups with the same number of elements isomorphic ? **(Meerut 1981)** **Ans. No.**

5. Show that the multiplicative group $G=\{1, \omega, \omega^2\}$ is isomorphic to the group G' of residue classes (mod 3) under addition of residue classes.

6. Show that the multiplicative group $G=\{1, -1, i, -i\}$ is isomorphic to the group $G'=\{0, 1, 2, 3\}$ with addition modulo 4 as composition. **(Rajasthan 69, Meerut 80)**

7. If f is an isomorphism of a group G onto a group G', show that f^{-1} is an isomorphism of G' onto G. **(Lucknow 68)**

8. Show that the product of two isomorphisms is also an isomorphism. **(Lucknow 68)**

9. Let G be any group and a be any fixed element in G. Define a mapping $f : G \to G$ by the formula

$$f(x)=axa^{-1}, \ \forall \ x \in G.$$

Prove that f is an isomorphism of G onto itself.

(I.A.S. 1970; Poona 73)

10. Prove that the order of an element of a group is unaltered by an isomorphism.

11. Show that the mapping $x \to x^{-1}$ of G onto G is an isomorphism if and only if G is abelian, x being any element of the group G. **(Poona 1973)**

12. Prove that the additive group of complex numbers $a+ib$ (a, b integers) is isomorphic to the multiplicative group of rational numbers of the form $2^a \, 3^b$ (a, b integers).

13. Let n be an integer greater than 1. Prove that

$$G=\left\{z_k : z_k=\cos \frac{2k\pi}{n}+i \sin \frac{2k\pi}{n} \ ; \ 0 \leqslant k \leqslant n-1 \right\}$$

is a group with respect to the multiplication of complex numbers. Show that G is isomorphic to $Z/(n)$ —the additive group of integers modulo n. **(Meerut 73)**

14. Show that the multiplicative group G of n n^{th} roots of unity is isomorphic to the group

$$G'=(\{0, 1, 2, \ldots, n-1\}, +_n).$$

15. Show that the group G of four transformations f_1, f_2, f_3, f_4 defined by $f_1(z)=z$, $f_2(z)=-z$, $f_3(z)=1/z$, $f_4(z)=-1/z$, with composite composition is isomorphic to the permutation group G' of degree four consisting of the four permutations

$$I,\ (ab),\ (cd),\ (ab)\ (cd).$$

16. Show that real matrices of the type $\begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix}$, where $a \neq 0$ form a multiplicative group which is isomorphic to the group of real non-zero numbers under multiplication.

17. Show that the multiplicative group of all matrices $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, where a and b are real numbers (not both equal to zero), is isomorphic to the group of non-zero complex numbers for multiplication. **(Meerut 72)**

§ 22. Complexes and subgroups of a group.

Suppose G is a group and the group composition has been denoted multiplicatively.

Any non-empty subset H of a group G is called a complex of the group G. **(Nagarjuna 1978)**

Let, now, H be any non-empty subset of a group G. Suppose H is closed with respect to the composition in G *i.e.*,

$$a \in H,\ b \in H \Rightarrow ab \in H.$$

Then we say that the complex H is **stable** for the composition in G and that the composition in G has **induced** a composition in H. This composition in H is called the **induced composition**. If for this induced composition H itself is a group, then H is called a subgroup of the group G.

Subgroup. Definition. **(Madras 1974; Usmania 72; Meerut 77, 82; Rajasthan 77; Garhwal 76)**

A non-empty subset H of a group G is said to be a subgroup of G if the composition in G is also a composition in H and for this composition H itself is a group.

Every subgroup of G is a complex of G but every complex is not always a subgroup.

Now every set is a subset of itself. Therefore if G is a group, then G itself is a subgroup of G. Also if e is the identity of G, then the subset of G containing only one element *i.e.*, e is also a subgroup of G. These two are subgroups of any group. They are called **trivial** or **improper** subgroups. A subgroup other than these two is called a **proper subgroup**.

Theorem. *Prove that*

(*i*) *The identity of a subgroup is the same as that of the group.*

(*ii*) *The inverse of any element of a subgroup is the same as the inverse of the same regarded as an element of the group.*

(*iii*) *The order of any element of a subgroup is the same as the order of the element regarded as a member of the group.*

Proof. Let H be a subgroup of the group G.

(i) Let e and e' be the identities of G and H respectively.

Now $a \in H \Rightarrow e'a = a.$ [$\because$ e' is identity of H]

Also $a \in H \Rightarrow a \in G \Rightarrow ea = a$ [$\because$ e is identity of G]

$\therefore$ in G we have $e'a = ea$

$\Rightarrow e' = e$ [by right cancellation law in G]

(ii) Let e be the identity of G as well as of H.

Let $a \in H$. Suppose b is the inverse of a in H and c is the inverse of a in G. Then we have $ba = e$ and $ca = e$.

$\therefore$ in G we have $ba = ca \Rightarrow b = c$.

(iii) It can also be easily proved.

Some Examples of Subgroups.

Example 1. The multiplicative group $\{1, -1\}$ is a subgroup of the multiplicative group $\{1, -1, i, -i\}$.

Example 2. The additive group of even integers is a subgroup of the additive group of all integers.

Example 3. The multiplicative group of positive rational numbers is a subgroup of the multiplicative group of all non-zero rational numbers.

Example 4. The additive group of integers is a subgroup of the additive group of all rational numbers.

Example 5. The alternating group A_n of all even permutations of degree n is a subgroup of the **symmetric group** P_n of all permutations of degree n.

Example 6. The subset $\{I, (1\ 2\ 3), (1\ 3\ 2)\}$ is a subgroup of the symmetric group P_3. Other proper subgroups of P_3 are $\{I, (1\ 2)\}, \{I, (2\ 3)\}, \{I, (3\ 1)\}$.

§ 23. Algebra of complexes of a group Multiplication of complexes. *If H and K are two complexes of a group G, then*

$$HK = \{x \in G \mid x = hk,\ h \in H,\ k \in K\}.$$

Obviously $HK \subseteq G$. Thus HK is a complex of G consisting of the elements of G obtained on multiplying each member of H with each member of K.

Multiplication of complexes is associative *i.e., If H, K, L are any three complexes of G, then $(HK)\,L=H\,(KL)$.* **(Nagarjuna 1978)**

Proof. Let h, k, l be any arbitrary elements of H, K, L respectively so that $(hk)\,l \in (HK)\,L$.

But $\quad (hk)\,l=h\,(kl) \in H\,(KL)$.

$\therefore \quad (HK)\,L \subseteq H\,(KL)$.

Similarly we can show that $H\,(KL) \subseteq (HK)\,L$.

Hence $\quad (HK)\,L=H\,(KL)$.

Note. If we say that $HK=KH$, then it does not mean that we should have $hk=kh$ for all $h \in H$ and for all $k \in K$. What we require is that each element of the set HK should be present in KH and each element of KH should be present in HK.

Inverse of a complex. *Let H be any complex of G. Then we define $H^{-1}=\{h^{-1} : h \in H\}$ i.e., H^{-1} is the complex of G consisting of the inverses of the elements of H.*

Theorem 1. *If H and K are any two complexes of a group G, then $(HK)^{-1}=K^{-1}\,H^{-1}$.* **(Nagarjuna 1978)**

Proof. Let x be any arbitrary element of $(HK)^{-1}$. Then

$x=(hk)^{-1},\ h\in H,\ k\in K$

$=k^{-1}\,h^{-1} \in K^{-1}\,H^{-1} \qquad [\because\ k^{-1} \in K^{-1},\ h^{-1} \in H^{-1}]$

$\therefore\ (HK)^{-1} \subseteq K^{-1}\,H^{-1}$.

Again let y be any arbitrary element of $K^{-1}\,H^{-1}$.

Then $\ y=k^{-1}\,h^{-1},\ k \in K,\ h \in H$

$=(hk)^{-1} \in (HK)^{-1}. \qquad [\because\ hk \in HK]$

$\therefore\ K^{-1}\,H^{-1} \subseteq (HK)^{-1}$.

Hence $(HK)^{-1}=K^{-1}\,H^{-1}$.

Theorem 2 *If H is any subgroup of G then $H^{-1}=H$. Also show that the converse is not true.*

Proof. Let h^{-1} be any arbitrary element of H^{-1}. Then $h\in H$. Now H is a subgroup of G. Therefore $h \in H \Rightarrow h^{-1} \in H$. Thus $h^{-1}\in H^{-1} \Rightarrow h^{-1} \in H$. Therefore $H^{-1}\subseteq H$.

Again $\quad h \in H \Rightarrow h^{-1}\in H \qquad [\because\ H$ is itself a group]

$\Rightarrow (h^{-1})^{-1}\in H^{-1} \qquad$ [by def. of H^{-1}]

$\Rightarrow h \in H^{-1}$.

$\therefore\ H \subseteq H^{-1}$.

Hence $H^{-1}=H$.

If H is a complex of a group G and $H^{-1}=H$, then it is not necessary that H is a subgroup of G. For example $H=\{-1\}$ is a complex of the multiplicative group $G=\{-1, 1\}$. Also $H^{-1}=\{-1\}$

since -1 is the inverse of -1 in G. But $H=\{-1\}$ is not a subgroup of G. We have $(-1)(-1)=1 \notin H$. Thus H is not closed with respect to multiplication.

Theorem 3. *If H is any subgroup of a group G, then*

$$HH=H.$$

Proof. Let h_1h_2 be any element of HH where $h_1 \in H$, $h_2 \in H$. Since H is a subgroup of G, therefore

$$h_1, h_2 \in H \Rightarrow h_1h_2 \in H.$$

$$\therefore \quad HH \subseteq H.$$

Now let h be any element of H. Then we can write $h=he$ where e is the identity of G. Now $he \in HH$, since $h \in H$, $e \in H$.

Thus $\quad H \subseteq HH.$

Hence $\quad HH=H.$

§ 24. Criterion for a complex to be a subgroup.

Theorem 1. *A non-empty subset H of a group G is a subgroup of G if and only if*

(*i*) $a \in H, b \in H \Rightarrow ab \in H.$

(*ii*) $a \in H \Rightarrow a^{-1} \in H$ *where a^{-1} is the inverse of a in G.*

(Lucknow 1970; Allahabad 69)

Proof. The conditions are necessary. Suppose H is a subgroup of G. Then H must be closed with respect to multiplication *i.e.*, the composition in G. Therefore $a \in H, b \in H \Rightarrow ab \in H$.

Let $a \in H$ and let a^{-1} be the inverse of a in G. Then the inverse of a in H is also a^{-1}. Since H itself is a group, therefore each element of H must possess inverse. Therefore

$$a \in H \Rightarrow a^{-1} \in H.$$

The conditions are sufficient.

Since $a \in H$, $b \in H \Rightarrow ab \in H$, therefore H is closed with respect to multiplication.

Associativity. The elements of H are also the elements of G. The composition in G is associative. Therefore the same composition must also be associative in H.

Existence of Identity. The identity of the subgroup is the same as the identity of the group.

Now $\quad a \in H \Rightarrow a^{-1} \in H.$ $\quad$ [From the given condition (ii)]

Further $\quad a \in H, a^{-1} \in H \Rightarrow aa^{-1} \in H$

[From the given condition (i)]

$$\Rightarrow e \in H.$$

$\therefore$ The identity e is an element of H.

Existence of inverse. Since $a \in H \Rightarrow a^{-1} \in H$, therefore each element of H possesses inverse. Hence H itself is a group for the composition in G. So H is a subgroup of G.

Theorem 3. (An important characteristic property of a subgroup). *A necessary and sufficient condition for a non-empty subset H of a group G to be a subgroup is that*

$a \in H$, $b \in H \Rightarrow ab^{-1} \in H$ *where b^{-1} is the inverse of b in G.*

(I. A. S. 1972 ; Nagarjuna 79, 80 ; Lucknow 80 ; Marathwada 70 ; Osmania 72 ; Meerut 81, 84 ; Gorakhpur 70 ; Allahabad 79 ; Punjab 70 ; Rohilkhand 80 ; Garhwal 76 ; Kumayon 78 ; Poona 73)

Proof. The condition is necessary. Suppose H is a subgroup of G. Let $a \in H$, $b \in H$. Now each element of H must possess inverse because H itself is a group.

$\therefore \quad b \in H \Rightarrow b^{-1} \in H.$

Further H must be closed with respect to multiplication *i.e.*, the composition in G. Therefore $a \in H$, $b^{-1} \in H \Rightarrow ab^{-1} \in H$.

The condition is sufficient. Now it is given that

$$a \in H, b \in H \Rightarrow ab^{-1} \in H.$$

We are to prove that H is a sub-group of G.

Existence of Identity. We have

$a \in H, a \in H \Rightarrow aa^{-1} \in H$ [by the given condition]

$\Rightarrow e \in H.$

Thus the identity e is an element of H.

Existence of Inverse. Let a be any element of H. Then by the given condition, we have $e \in H, a \in H \Rightarrow ea^{-1} \in H$

$\Rightarrow a^{-1} \in H.$

Thus each element of H possesses inverse.

Closure Property. Let $a, b \in H$. Then as shown above $b \in H \Rightarrow b^{-1} \in H$. Therefore applying the given condition, we have $a \in H, b^{-1} \in H \Rightarrow a\,(b^{-1})^{-1} \in H \Rightarrow ab \in H$.

Associativity. The elements of H are also the elements of G. The composition in G is associative. Therefore it must also be associative in H.

Hence H itself is a group for the composition in G. Therefore H is a subgroup of G.

Note. Had we denoted the composition in G additively, the statement of the above theorem would have been

$$a \in H, b \in H \Rightarrow a-b \in H.$$

An Important Remark. The above theorem gives us a very important characterization of a subgroup. Whenever we are required to show that a non-empty subset H of a group G is a subgroup of G, we should make use of this theorem. Thus we should take any two arbitrary elements $a, b \in H$ and we should try to show that $ab^{-1} \in H$. If we are able to show that $a, b \in H \Rightarrow ab^{-1} \in H$, it is sufficient to conclude that H is a subgroup of G.

Cor. 1. *A necessary and sufficient condition for a non-empty subset H of a group G to be a subgroup is that* $HH^{-1} \subseteq H$.

Proof. The condition is necessary. It is given that H is a subgroup of G. Let ab^{-1} be any arbitrary element of HH^{-1}. Then $a \in H, b \in H$.

Since H itself is a group, therefore $b \in H \Rightarrow b^{-1} \in H$.

Thus $a \in H, b^{-1} \in H \Rightarrow ab^{-1} \in H$, by closure property.

$\therefore \quad ab^{-1} \in HH^{-1} \Rightarrow ab^{-1} \in H.$

Hence $\qquad HH^{-1} \subseteq H.$

The condition is sufficient. It is given that $HH^{-1} \subseteq H$.

Let $a, b \in H$. Then $ab^{-1} \in HH^{-1}$. Since $HH^{-1} \subseteq H$, therefore $ab^{-1} \in HH^{-1} \Rightarrow ab^{-1} \in H$. Thus $a \in H, b \in H \Rightarrow ab^{-1} \in H$. Hence H is a subgroup of G.

Cor. 2. *A necessary and sufficient condition for a non-empty subset H of a group G to be a subgroup is that* $HH^{-1} = H$.

Proof. The condition is necessary Suppose H is a sub-group of G. Then by Cor. 1, $HH^{-1} \subseteq H$.

Now H is a sub-group of G. Therefore $e \in H$. If h is any arbitrary element of H, then

$$h = he = he^{-1} \in HH^{-1} \qquad [\because \; h \in H, e^{-1} \in H^{-1}]$$

$\therefore \quad H \subseteq HH^{-1}.$

Hence $\qquad HH^{-1} = H.$

The condition is sufficient. It is given that $HH^{-1} = H$.

$\therefore \quad HH^{-1} \subseteq H.$

Hence by Cor. 1, H is a sub-group of G.

Theorem 3. (Criterion in the case of finite complexes).

A necessary and sufficient condition for a non-empty finite subset H of a group G (the composition in G being denoted multiplicatively) to be a sub-group is that H must be closed with respect to multiplication i.e., $a \in H, b \in H \Rightarrow ab \in H$.

(Kerala 1970; Madurai 78; Meerut 78, 82; Rajasthan 77; Sagar 77; Kanpur 80; Banaras 70)

Proof. The condition is necessary. Suppose H is a sub-group of G. Then H must be closed with respect to multiplication *i.e.*, the composition in G. Therefore $a \in H, b \in H \Rightarrow ab \in H$. Hence the condition is necessary.

The condition is sufficient. It is given that H is closed with respect to multiplication *i.e.*, $a \in H$, $b \in H \Rightarrow ab \in H$.

Let a be any element of H. Then by the given condition $a^2 = aa \in H$, $a^3 = aa^2 \in H$, $a^4 = aa^3 \in H$. Proceeding in this way we get $a^m \in H$ where m is any positive integer. Thus the infinite collection of elements $a, a^2, a^3, \ldots, a^n, \ldots$ all belong to H. But H is a finite subset of G. Therefore there must be repititions in this collection of elements. If they are all distinct, then H will not be a finite set. Therefore for some positive integers r and s with $r > s$, we must have

$$a^r = a^s$$

$$\Rightarrow a^r a^{-s} = a^s a^{-s} \qquad [\because a^s \in G \Rightarrow (a^s)^{-1} \text{ i.e., } a^{-s} \in G]$$

$$\Rightarrow a^{r-s} = a^0 = e, \text{ where } e \text{ is the identity of } G.$$

$\because$ $r-s$ is a positive integer, therefore $a^{r-s} = e \in H$.

Therefore the identity e *i.e.*, a^0 is also an element of H.

Now $r - s \geqslant 1$. Therefore $r - s - 1 \geqslant 0$.

We have $a^{r-s-1} a = a^{r-s} = e = a a^{r-s-1}$.

$\therefore$ by the definition of inverse, $a^{-1} = a^{r-s-1} \in H$. Thus each element of H possesses inverse.

Finally the elements of H are also the elements of G. Therefore the composition in H must be associative.

Hence H is a subgroup of G.

Corollary. *A finite non-empty subset H of a group G is a subgroup of G iff HH=H.*

An Important Remark. The criterion given in the above theorem 3 is valid only for finite subsets of a group G. It is not valid for infinite subsets of an infinite group G as is clear from the following examples :

Example 1. Let G be the additive group of all integers and H be the subset of G consisting of all positive integers. Obviously H is closed with respect to addition *i.e.*, the composition in G. But H is not a subgroup of G since the identity $0 \notin H$.

Example 2. Let $G = \{\ldots\ldots 2^{-3}, 2^{-2}, 2^{-1}, 1, 2, 2^2, 2^3, \ldots\ldots\}$ be the multiplicative group consisting of all integral powers of 2. Let $H = \{1, 2, 2^2, 2^3, \ldots\}$. Then $H \subseteq G$ and H is closed with respect to multiplication. But H is not a subgroup of G since the inverse of 2 *i.e.*, 2^{-1} does not belong to H.

Example 3. Let G be the multiplicative group of all non-zero rational numbers. Then $H=\{1, -1\}$ is a finite subset of G. Also H is closed with respect to multiplication. Therefore H is a subgroup of G.

Example 4. Let P_n be the symmetric group of degree n *i.e.*, the elements of P_n are permutations of degree n. If A_n is the set of all even permutations of degree n, then $A_n \subseteq P_n$ and A_n is closed with respect to multiplication of permutations. We remember that the product of two even permutations is also an even permutation. Therefore A_n is a subgroup of P_n. Hence A_n is itself a group with respect to multiplication of permutations. A_n is called **alternating group of degree n or order $\frac{n!}{2}$**

§ 25. Criterion for the product of two subgroups to be a subgroup.

Theorem 1. *If H, K are two subgroups of a group G, then HK is a subgroup of G, iff $HK=KH$.* **(Vikram 1976; Madras 83; Nagarjuna 78; Raj. 78; Andhra 75; Meerut 78, 79)**

Proof. Let H and K be any two subgroups of a group G. Let $HK=KH$. In order to prove that HK is a subgroup of G it is sufficient to prove that $(HK)(HK)^{-1}=HK$.

We have $(HK)(HK)^{-1}=(HK)(K^{-1}H^{-1})=H(KK^{-1})H^{-1}$

$$\begin{aligned}&=(HK)H^{-1} \quad [\because \; K \text{ is a subgroup} \Rightarrow KK^{-1}=K]\\&=(KH)H^{-1} \quad [\because \; HK=KH]\\&=K(HH^{-1})\\&=KH \quad [\because \; H \text{ is a subgroup} \Rightarrow HH^{-1}=H]\\&=HK.\end{aligned}$$

$\therefore \quad HK=KH \Rightarrow HK$ is a subgroup of G.

Conversely suppose that HK is a subgroup.

Then $(HK)^{-1}=HK$

$$\begin{aligned}&\Rightarrow K^{-1}H^{-1}=HK\\&\Rightarrow KH=HK. \qquad [\because \; K \text{ is subgroup} \Rightarrow K^{-1}=K \text{ and similarly } H^{-1}=H]\end{aligned}$$

Hence the result.

Corollary. *If H, K are subgroups of an abelian group G, then HK is a subgroup of G.*

Proof. We know that if H, K are two subgroups of a group G, then HK is a subgroup of G if and only if $HK=KH$.

Since the given group G is here abelian, therefore we have $HK=KH$. Hence HK is a subgroup of G.

§ 26. Intersection of Subgroups.

Theorem 1. *If H_1 and H_2 are two subgroups of a group G, then $H_1 \cap H_2$ is also a subgroup of G.*

(Meerut 1970; Rajasthan 75; Kumayon 77; Poona 73; Jabalpur 70)

Proof. Let H_1 and H_2 be any two subgroups of G. Then $H_1 \cap H_2 \neq \varnothing$, since at least the identity element e is common to both H_1 and H_2.

In order to prove that $H_1 \cap H_2$ is a subgroup it is sufficient to prove that $a \in H_1 \cap H_2$, $b \in H_1 \cap H_2 \Rightarrow ab^{-1} \in H_1 \cap H_2$.

Now $a \in H_1 \cap H_2 \Rightarrow a \in H_1$ and $a \in H_2$,

$b \in H_1 \cap H_2 \Rightarrow b \in H_1$ and $b \in H_2$.

But H_1, H_2 are subgroups. Therefore

$a \in H_1$, $b \in H_1 \Rightarrow ab^{-1} \in H_1$,

$a \in H_2$, $b \in H_2 \Rightarrow ab^{-1} \in H_2$.

Finally, $ab^{-1} \in H_1$, $ab^{-1} \in H_2 \Rightarrow ab^{-1} \in H_1 \cap H_2$.

Thus we have shown that

$a \in H_1 \cap H_2$, $b \in H_1 \cap H_2 \Rightarrow ab^{-1} \in H_1 \cap H_2$.

Hence $H_1 \cap H_2$ is a subgroup of G.

Theorem 2. *Arbitrary intersection of subgroups i.e., the intersection of any family of subgroups of a group is a subgroup.*

(Kerala 1970; Nagarjuna 78)

Proof. Let G be a group and let $\{H_t : t \in T\}$ be any family of subgroups of G. Here T is an index set and is such that $\forall\ t \in T$, H_t is a subgroup of G.

Let $H = \bigcap_{t \in T} H_t = \{x \in G : x \in H_t \ \forall\ t \in T\}$

be the intersection of this family of subgroups of G. Then to prove that H is also a subgroup of G.

Obviously $H \neq \varnothing$, since at least the identity element e is in $H_t \ \forall\ t \in T$.

Now let a, b be any two elements of H. Then

$$a \in \bigcap_{t \in T} H_t \Rightarrow a \in H_t \ \forall\ t \in T,$$

and
$$b \in \bigcap_{t \in T} H_t \Rightarrow b \in H_t \ \forall\ t \in T.$$

But $\forall\ t \in T$, H_t is a subgroup of G. Therefore

$$a \in H_t,\ b \in H_t \Rightarrow ab^{-1} \in H_t \ \forall\ t \in T.$$

Consequently $ab^{-1} \in \bigcap_{t \in T} H_t$.

Thus we have shown that $a, b \in \bigcap_{t \in T} H_t \Rightarrow ab^{-1} \in \bigcap_{t \in T} H_t$.

Therefore $\bigcap_{t \in T} H_t$ is a subgroup of G.

Note 1. $H_1 \cap H_2$ is the largest subset of G which is contained in H_1 as well as in H_2. Therefore $H_1 \cap H_2$ is the largest subgroup of G contained in H_1 and H_2. By largest we mean that it is contained in H_1 and H_2 and contains every subgroup of G contained in both H_1 and H_2.

Note 2. *The union of two subgroups is not necessarily a subgroup.* **(Allahabad 1979)**

For example, let G be the additive group of integers.

Then $H_1 = \{\ldots, -6, -4, -2, 0, 2, 4, 6, \ldots\}$ and

$H_2 = \{\ldots, -12, -9, -6, -3, 0, 3, 6, 9, 12, \ldots\}$

are both subgroups of G.

We have $H_1 \cup H_2 = \{\ldots, -4, -3, -2, 0, 2, 3, 4, 6, \ldots\}$.

Obviously $H_1 \cup H_2$ is not closed with respect to addition as $2 \in H_1 \cup H_2$, $3 \in H_1 \cup H_2$ but $2+3$ *i.e.*, $5 \notin H_1 \cup H_2$. Therefore $H_1 \cup H_2$ is not a subgroup of G.

However $H_1 \cap H_2 = \{\ldots, -18, -12, -6, 0, 6, 12, 18, \ldots\}$ is a subgroup of G.

If we take the subgroup $H_3 = \{\ldots, -8, -4, 0, 4, 8, \ldots\}$ of G, then $H_1 \cup H_3 = H_1$ and H_1 is a subgroup of G. We shall prove in one of the following examples that the union of two subgroups is a subgroup iff one is contained in the other.

Solved Examples

Ex. 1. *Let G be the additive group of integers. Then prove that the set of all multiples of integers by a fixed integer m is a subgroup of G.*

Solution. We have $G = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ is the additive group of integers. Let m be any fixed integer. Let $H = \{\ldots, -3m, -2m, -m, 0, m, 2m, 3m, \ldots\}$. Then $H \subseteq G$. To prove that H is a subgroup of G. Let $a = rm$ and $b = sm$ be any two elements of H where r and s are some integers. The inverse of sm in G is $(-s)\,m$ *i.e.*, $-b = (-s)\,m$.

We have $a - b = rm + (-s)\,m = (r-s)\,m \in H$ since $r-s$ is also some integer. Thus $a \in H, b \in H \Rightarrow a - b \in H$. Hence H is a subgroup of G.

Ex. 2. *Let a be an element of a group G. The set $H = \{a^n : n \in \mathbf{I}\}$ of all integral powers of a is a subgroup of G.*

Solution. We have $a \in G$. To prove that

$H=\{\ldots, a^{-3}, a^{-2}, a^{-1}, a^{0}, a^{1}, a^{2}, a^{3}, \ldots\}$ is a subgroup of G.

Let a^r, a^s be any two elements of H, where r and s are some integers. The inverse of a^s in G is a^{-s}. Now

$a^r (a^s)^{-1}=a^r a^{-s}=a^{r-s} \in H$ since $r-s$ is also some integer. Therefore H is a subgroup of G.

Note. If G is a group and $a \in G$, then the subgroup $H=\{a^n : n \in I\}$ of G is called the subgroup of G *generated* by a.

Illustration. Let G be the multiplicative group of positive rational numbers. We have $3 \in G$. Obviously

$H=\{\ldots, 3^{-3}, 3^{-2}, 3^{-1}, 3^{0}, 3^{1}, 3^{2}, 3^{3}, \ldots\}$ is a subgroup of G.

Ex. 3. *Let G be the set of all ordered pairs (a, b) of real numbers for which $a \neq 0$. Let a binary operation $\times$ on G be defined by the formula*

$$(a, b) \times (c, d)=(ac, bc+d).$$

Show that $(G, \times)$ is a non-abelian group.

Does the subset H of all those elements of G which are of th form $(1, b)$ form a subgroup of G? **(Madras 1974; Rajasthan 69)**

Solution. For the first part of the question see example 7 on page 65. The inverse of the element (a, b) of G has been found to be $\left(\frac{1}{a}, -\frac{b}{a}\right)$. Thus $(a, b)^{-1}=\left(\frac{1}{a}, -\frac{b}{a}\right)$.

Now we are to see whether H is a subgroup of G or not. Obviously H is a non-empty subset of G. Let $(1, b)$ and $(1, c)$ be any two elements of H. Then

$$(1, b) \times (1, c)^{-1}=(1, b) \times \left(\frac{1}{1}, -\frac{c}{1}\right)=(1, b) \times (1, -c)=(1, b-c)$$

[by def. of the operation $\times$ on G]

But $(1, b-c)$ is definitely an element of H. Thus

$$(1, b), (1, c) \in H \Rightarrow (1, b) \times (1, c)^{-1} \in H.$$

Hence H is a subgroup of G.

Ex. 4. *Let H be a subgroup of a group G and define*

$$T=\{x \in G : xH=Hx\}.$$

Prove that T is a subgroup of G.

Solution. Let $x_1, x_2 \in T$. Then $x_1H=Hx_1$, $x_2H=Hx_2$.

First we show that $x_2^{-1} \in T$.

We have $x_2H=Hx_2 \Rightarrow x_2^{-1} (x_2H) x_2^{-1}=x_2^{-1} (Hx_2) x_2^{-1}$

$$\Rightarrow Hx_2^{-1}=x_2^{-1} H \Rightarrow x_2^{-1} \in T.$$

Now we shall show that $x_1x_2^{-1} \in T$.

We have $(x_1x_2^{-1})\, H = x_1\, (x_2^{-1}\, H) = x_1\, (Hx_2^{-1})$

$$= (x_1H)\, x_2^{-1} = (Hx_1)\, x_2^{-1} = H\, (x_1x_2^{-1}).$$

$\therefore \quad x_1x_2^{-1} \in T.$

Thus $x_1,\, x_2 \in T \Rightarrow x_1x_2^{-1} \in T$. Hence T is a subgroup of G.

Ex. 5. *Prove that those elements of a group G which commute with the square of a given element b of G form a subgroup H of G and those which commute with b itself form a subgroup of H.*

(Poona 1973)

Solution. Let $H = \{x \in G : xb^2 = b^2x\}$. Then to prove that H is a subgroup of G. We see that H is not empty because $eb^2 = b^2 = b^2e \Rightarrow e \in H$.

Now let $x_1,\, x_2 \in H$. Then $x_1b^2 = b^2x_1$ and $x_2b^2 = b^2x_2$.

First we shall show that $x_2^{-1} \in H$. We have

$$x_2b^2 = b^2x_2 \Rightarrow x_2^{-1}\, (x_2b^2)\, x_2^{-1} = x_2^{-1}(b^2x_2)\, x_2^{-1}$$

$$\Rightarrow b^2x_2^{-1} = x_2^{-1}b^2 \Rightarrow x_2^{-1} \in H.$$

Now we shall show that $x_1x_2^{-1} \in H$.

We have $x_1x_2^{-1}\, b^2 = x_1b^2x_2^{-1}$ $\quad [\because\ b^2x_2^{-1} = x_2^{-1}b^2]$

$= b^2x_1x_2^{-1}$ $\quad [\because\ x_1b^2 = b^2x_1]$

$\therefore \quad x_1x_2^{-1} \in H.$

Thus $x_1,\, x_2 \in H \Rightarrow x_1x_2^{-1} \in H \Rightarrow H$ is a subgroup of G.

Let $N = \{y \in G : yb = by\}$. We have $yb = by \Rightarrow (yb)\, b = (by)\, b$ $\Rightarrow yb^2 = b\, (yb) \Rightarrow yb^2 = b\, (by) \Rightarrow yb^2 = b^2y$.

Thus $y \in N \Rightarrow y \in H$. Therefore $N \subseteq H$.

Now to prove that N is a subgroup of H. Obviously N is not empty since at least $e \in N$.

Let $y_1,\, y_2 \in N$. Then $y_1b = by_1$ and $y_2b = by_2$.

We have $y_2b = by_2 \Rightarrow y_2^{-1}\, (y_2b)\, y_2^{-1} = y_2^{-1}\, (by_2)\, y_2^{-1}$

$$\Rightarrow by_2^{-1} = y_2^{-1}\, b.$$

Now $y_1\, y_2^{-1}\, b = y_1\, by_2^{-1} = by_1y_2^{-1}$.

$\therefore \quad y_1y_2^{-1} \in N.$

Thus $y_1,\, y_2 \in N \Rightarrow y_1y_2^{-1} \in N$. Hence N is a subgroup of H.

Ex. 6. *Show that the union of two subgroups is a subgroup if and only if one is contained in the other.*

(Rajasthan 1976; Allahabad 79; Madras 83)

Solution. Suppose H_1 and H_2 are two subgroups of a group G. Let $H_1 \subseteq H_2$ or $H_2 \subseteq H_1$. Then $H_1 \cup H_2 = H_2$ or H_1. But $H_1,\, H_2$ are subgroups and therefore $H_1 \cup H_2$ is also a subgroup.

Conversely suppose $H_1 \cup H_2$ is a subgroup. To prove that $H_1 \subseteq H_2$ or $H_2 \subseteq H_1$.

Let us assume that H_1 is not a subset of H_2 and H_2 is also not a subset of H_1.

Now H_1 is not a subset of $H_2 \Rightarrow \exists\, a \in H_1$ and $a \notin H_2$ ...(1)
and H_2 is not a subset of $H_1 \Rightarrow \exists\, b \in H_2$ and $b \notin H_1$. ...(2)

From (1) and (2), we have $a \in H_1 \cup H_2$ and $b \in H_1 \cup H_2$.

Since $H_1 \cup H_2$ is a subgroup, therefore $ab=c$ (say) is also an element of $H_1 \cup H_2$.

But $ab=c \in H_1 \cup H_2 \Rightarrow ab=c \in H_1$ or H_2.

Suppose $ab=c \in H_1$.

Then $b=a^{-1}c \in H_1$ [$\because$ H_1 is a subgroup, therefore $a \in H_1 \Rightarrow a^{-1} \in H_1$]

But from (2), we have $b \notin H_1$. Thus we get a contradiction.

Again suppose $ab=c \in H_2$.

Then $a=cb^{-1} \in H_2$. [$\because$ H_2 is a subgroup, therefore $b \in H_2 \Rightarrow b^{-1} \in H_2$]

But from (1), we have $a \notin H_2$. Thus here also we get a contradiction.

Hence either $H_1 \subseteq H_2$ or $H_2 \subseteq H_1$.

Ex. 7. *Let G be the multiplicative group of all positive real numbers and R the additive group of all real numbers. Is G a subgroup of R ?* **(Gujrat 1970)**

Solution. There is no doubt that the set G of all positive real numbers is a subset of the set R of all real numbers. But the group G is not a subgroup of the group R. The reason is that the composition in G is different from the composition in R.

Ex. 8. (*i*) *Can an abelian group have a non-abelian sub-group ?* **(Nagarjuna 1979, 80)**

(*ii*) *Can a non-abelian group have an abelian sub-group ?* **(Nagarjuna 1979)**

(*iii*) *Can a non-abelian group have a non-abelian sub-group ?* **(Nagarjuna 1980)**

Solution. (i) Every sub-group of an abelian group is abelian. If G is an abelian group and H is a sub-group of G, then the operation on H is commutative because it is already commutative in G and H is a subset of G. Hence an abelian group cannot have a non-abelian sub-group.

(ii) A non-abelian group can have an abelian sub-group. For example the symmetric group P_3 of permutations of degree 3 is non-abelian while its sub-group A_3 is abelian.

(iii) A non-abelian group can have a non-abelian sub-group. For example P_4 is a non-abelian group and its sub-group A_4 is also non-abelian.

Exercises

1. (i) Define a subgroup. Give examples. **(Meerut 1975)**

(ii) What is the difference between a complex and a sub-group of a group ? **(Meerut 1979)**

2. Show that a necessary and sufficient condition for a non-empty subset H of a finite group G to be a sub-group is that

$$a \in H,\ b \in H \Rightarrow ab \in H.$$ **(Banaras 1971)**

3. (*i*) Define alternating group. Show that the alternating group A_n is a subgroup of the symmetric group S_n of the permutations over n objects. Write down all the proper subgroups of S_3. **(Gorakhpur 1970)**

Ans. $\{I, (ab)\}, \{I, (bc)\}, \{I, (ca)\}, \{I, (abc), (acb)\}$.

(*ii*) Give an example of a non-abelian group G which has the property that every proper subgroup of G is abelian.

(Meerut 1980)

4. H is a finite non-empty subset of a group G and is closed with respect to the group operation. Prove that H is a subgroup of G. Show also, by means of an example, that the above statement is not necessarily true when H is an infinite subset of G.

(Kerala 1972; Marathwada 72)

5. Verify the following statements for being true or false. In case a statement is false, write the corresponding correct statement.

(i) A non-empty subset H of a group G, which is closed under the binary composition in G is a subgroup of G.

(ii) If G is a group and H is a non-empty subset of G, then H will be a subgroup of G if $H^2=H$. **(Meerut 1976)**

Ans. (i) False ; (ii) False.

6. If G is a group, the centre of G, Z is defined by

$$Z=\{z \in G : zx=xz \ \forall \ x \in G\}.$$

Prove that Z is a subgroup of G.

Or

Show that the elements in a group G which commute with every element of G form a sub-group of G. **(Meerut 1981)**

7. If $a \in G$ we define $N(a)=\{x \in G : xa=ax\}$.

Show that $N(a)$ is a subgroup of G.

(Punjab 1970; Meerut 76, 78, 79, 81; Delhi 70)

8. Let G be a group, H a subgroup of G. Let for $x \in G$,

$$xHx^{-1}=\{xhx^{-1} : h \in H\}.$$

Prove that xHx^{-1} is a subgroup of G. **(Meerut 1979)**

9. Show that the elements of finite order in any commutative group G form a subgroup of G.

10. Show by means of examples that the union of two subgroups may or may not be a subgroup. **(Rajasthan 1976)**

11. Show that the integral multiples of 5 form a subgroup of the additive group of integers. **(Meerut 1973)**

12. Let A and B be subgroups of a group G and let AB be the subset of G consisting of all elements of the form ab, where a is in A and b is in B. Then

(i) Show by considering the subgroups of the group of permutations of three elements, or otherwise, that AB need not be a subgroup of G.

(ii) Show that AB is a subgroup of G if and only if $AB=BA$.

(Gujrat 1971)

13. Show that the 24 permutations on 4 symbols form a group with respect to permutation multiplication. Write down three proper subgroups of this group. **(Kanpur 1970)**

14. Show that all those elements of an abelian group G which satisfy the relation $a^2=e$ constitute a subgroup of G.

15. Show that a group can never be expressed as the union of two of its proper subgroups. **(Poona 1973)**

16. Let the mapping π_{ab}, for a, b real numbers, map the reals into the reals by the rule, $\pi_{ab}(x)=ax+b$.

Let $G=\{\pi_{ab} : a \neq 0\}$. Prove that G is a group under the composition of product of mappings. Find the formula for $\pi_{ab}\,\pi_{cd}$.

Let $H=\{\pi_{ab} \in G : a \text{ is rational}\}$. Show that H is a subgroup of G. **(I.A.S. 1973)**

17. Consider the set

$$S=\left\{\frac{a}{2^n} : a \text{ is any integer, } n \text{ is a fixed integer} \geqslant 0\right\}.$$

Show that S is a subgroup of the additive group of rational numbers and that S contains the additive group of integers.

(Meerut 1974)

18. Let A be any non-empty set and $x \in A$ be a fixed element. Let $T_x = \{\sigma \mid \sigma : A \to A$ is a permutation and $\sigma(x) = x\}$ Show that T_x is a subgroup of the group of all permutations of A.

(Meerut 1980)

§ 27. **Cosets.** We shall now introduce the very important concept of **right and left cosets of any subgroup.** These are also known as **residue classes modulo the subgroup.** Cosets of a subgroup are only special types of complexes.

Definition. *Suppose G is a group and H is any subgroup of G. Let a be any element of G. Then the set* $Ha = \{ha : h \in H\}$ *is called a right coset of H in G generated by a. Similarly the set* $aH = \{ah : h \in H\}$ *is called a left coset of H in G generated by a.*

(Meerut 1972; Kumayon 77)

Obviously Ha and aH are both subsets of G.

If e is the identity element of G, then $He = H = eH$. Therefore *H itself is a right as well as a left coset.*

Since H is a subgroup of G, therefore $e \in H$. So if Ha is a right coset of H in G, then ea is an element of Ha. Thus we see that $a \in Ha$. Therefore if Ha is any right coset, then at least the element $a \in Ha$. *Consequently no right coset can be empty.* Similarly a is an element of the left coset aH. *Therefore no left coset can be empty.*

If the group G is abelian, then we have $ah = ha \ \forall \ h \in H$. Therefore the right coset Ha will be equal to the corresponding left coset aH. However if the group G is not abelian, then we may have $aH = Ha$ or $aH \neq Ha$.

Note. If the composition in the group G has been denoted additively, then the right coset of H in G generated by a is defined as

$$H + a = \{h + a : h \in H\}.$$

Similarly the left coset $a + H = \{a + h : h \in H\}$.

Example 1. **(Rajasthan 1978)** Let G be the additive group of integers *i.e.*, $G = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$.

Let H be the subgroup of G obtained on multiplying each element of G by 3. Then $H = \{\ldots, -9, -6, -3, 0, 3, 6, 9, \ldots\}$.

Since the group G is abelian any right coset will be equal to the corresponding left coset. Let us form the right cosets of H in G.

We have $0 \in G$ and

$$H = H+0 = \{\ldots, -9, -6, -3, 0, 3, 6, 9, \ldots\}.$$

Again $1 \in G$ and $H+1 = \{\ldots, -8, -5, -2, 1, 4, 7, 10, \ldots\}$.

Then $2 \in G$ and $H+2 = \{\ldots, -7, -4, -1, 2, 5, 8, 11, \ldots\}$.

We see that the right cosets H, $H+1$ and $H+2$ are all distinct and moreover these are disjoint *i.e.*, have no element common.

Now $3 \in G$ and $H+3 = \{\ldots, -6, -3, 0, 3, 6, 9, 12, \ldots\}$.

We see that $H+3 = H$ Also we observe that $3 \in H$.

Again $4 \in G$ and $H+4 = \{\ldots, -5, -2, 1, 4, 7, 10, 13, \ldots\}$.

We see that $H+4 = H+1$. Also we observe that $4 \in H+1$.

Similarly the right coset $H+5$ coincides with $H+2$, $H+6$ with H, $H+(-1)$ with $H+2$, $H+(-2)$ with $H+1$ and so on.

Thus we get only three distinct right cosets *i.e.*, H, $H+1$, $H+2$. Obviously $G = H \cup (H+1) \cup (H+2)$.

Example 2. *Let G be the group of all permutations of degree* 3 *on three symbols* 1, 2, 3. *Then the elements of G are the permutations* $f_1 = (1), f_2 = (12), f_3 = (23), f_4 = (31), f_5 = (123), f_6 = (132)$.

Let H be the subgroup of G consisting of the permutations f_1 and f_2. Then $H = \{f_1, f_2\}$. Let us form the right cosets of H in G

$Hf_1 = H$, since f_1 is the identity element of G.

$Hf_2 = \{f_1 f_2, f_2 f_2\} = \{f_2, f_1\} = H$.

$Hf_3 = \{f_1 f_3, f_2 f_3\} = \{f_3, f_6\}$. [Note that $f_2 f_3 = f_6$]

$Hf_4 = \{f_1 f_4, f_2 f_4\} = \{f_4, f_5\}$.

$Hf_5 = \{f_1 f_5, f_2 f_5\} = \{f_5, f_4\}$. [Note that $Hf_4 = Hf_5$]

$Hf_6 = \{f_1 f_6, f_2 f_6\} = \{f_6, f_3\} = Hf_3$.

Thus we get only three distinct right cosets *i.e.*, H, Hf_3, Hf_4. Obviously $G = H \cup (Hf_3) \cup (Hf_4)$.

Also note that the number of elements in each right coset is the same as in H.

The left coset $f_3 H = \{f_3 f_1, f_3 f_2\} = \{f_3, f_5\}$.

We see that $Hf_3 \neq f_3 H$.

Forming all the left cosets of H in G, we can see that we shall get only three distinct left cosets.

Note. Suppose in example 2, we take $H = \{f_1, f_5, f_6\}$ which is obviously a subgroup of G. Of course H is nothing but the alternating group A_3. There will be only two distinct right cosets of H in G and each of them will contain three elements. It will be interesting to check that in the case of this subgroup of G each right coset is equal to the corresponding left coset.

Theorem 1. *If H is any subgroup of G and $h \in H$, then $Hh=H=hH$.*

Proof. Let $h \in H$. Then to prove that $Hh=H$. Suppose h' is any arbitrary element of H. Then $h'h$ is an arbitrary element of Hh.

Since H is a subgroup, we have

$$h' \in H, h \in H \Rightarrow h'h \in H.$$

Thus every element of Hh is also an element of H. Hence

$$Hh \subseteq H.$$

$$\text{Again } h'=h'\,(h^{-1}h) \qquad [\because\ h^{-1}h=e]$$
$$=(h'h^{-1})h$$
$$\in Hh\ [\because\ h \in H \Rightarrow h^{-1} \in H \text{ and } h' \in H,\ h^{-1} \in H \Rightarrow h'h^{-1} \in H].$$

Thus every element h' of H is also an element of Hh. Hence $H \subseteq Hh$.

Finally $Hh \subseteq H$ and $H \subseteq Hh \Rightarrow Hh=H$.

Similarly we can prove that $hH=H$.

Note. If H is any subgroup of G and $a \in G$, then $a \in Ha$. Now suppose $Ha=H$. Then each element of Ha belongs to H. Therefore $a \in H$. Hence $Ha=H \Rightarrow a \in H$.

Similarly $aH=H \Rightarrow a \in H$.

Theorem 2. *If a, b are any two elements of a group G and H any subgroup of G, then*

$Ha=Hb \Leftrightarrow ab^{-1} \in H$ and $aH=bH \Leftrightarrow a^{-1}b \in H$. **(Raj. 78)**

Proof. Since a is an element of Ha, therefore

$$Ha=Hb \Rightarrow a \in Hb \Rightarrow ab^{-1} \in (Hb)\,b^{-1}$$
$$\Rightarrow ab^{-1} \in H\,(bb^{-1}) \Rightarrow ab^{-1} \in He \Rightarrow ab^{-1} \in H.$$

$$\text{Conversely, } ab^{-1} \in H \Rightarrow Hab^{-1}=H \quad [\because\ h \in H \Rightarrow Hh=H]$$
$$\Rightarrow Hab^{-1}b=Hb$$
$$\Rightarrow Hae=Hb \Rightarrow Ha=Hb.$$

Similarly we can prove that $aH=bH \Leftrightarrow a^{-1}b \in H$.

Note. H is a subgroup. Therefore if $ab^{-1} \in H$, then $(ab^{-1})^{-1}$ i.e., $(b^{-1})^{-1}a^{-1}$ i.e., $ba^{-1} \in H$. Similarly if $a^{-1}b \in H$, then also $b^{-1}a \in H$.

Theorem 3. *If a, b are any two elements of a group G and H any subgroup of G, then*

$$a \in Hb \Leftrightarrow Ha=Hb \text{ and } a \in bH \Leftrightarrow aH=bH.$$

Proof. We have

$$\begin{aligned} a \in Hb &\Rightarrow ab^{-1} \in Hbb^{-1} \Rightarrow ab^{-1} \in He \\ &\Rightarrow ab^{-1} \in H \Rightarrow Hab^{-1}=H \\ &\Rightarrow Hab^{-1}\, b=Hb \Rightarrow Hae=Hb \Rightarrow Ha=Hb. \end{aligned}$$

Conversely, let $Ha=Hb$. Since $a \in Ha$, therefore $a \in Hb$.

Similarly we can prove that $a \in bH \Leftrightarrow aH=bH$.

Theorem 4. *Any two right (left) cosets of a subgroup are either disjoint or identical.* (Sagar 1966)

Proof. Suppose H is a subgroup of a group G and let Ha and Hb be two right cosets of H in G. Suppose Ha and Hb are not disjoint. Then there exists at least one element, say, c such that $c \in Ha$ and $c \in Hb$. Let $c=h_1a$ and $c=h_2b$, where $h_1, h_2 \in H$.

Then $h_1a=h_2b$

or $h_1^{-1}\, h_1a=h_1^{-1}\, h_2b$

or $ea=(h_1^{-1}\, h_2)\, b$

or $a=(h_1^{-1}\, h_2)\, b.$

Since H is a subgroup, therefore $h_1^{-1}\, h_2 \in H$. Let $h_1^{-1}\, h_2=h_3$. Then $a=h_3b$.

Now $$Ha=Hh_3b=(Hh_3)\, b =Hb. \qquad [\because\ h_3 \in H \Rightarrow Hh_3=H]$$

Therefore the two right cosets are identical if they are not disjoint. Thus either $Ha \cap Hb=\varnothing$ or $Ha=Hb$.

Similarly we can prove that either $aH \cap bH=\varnothing$ or $aH=bH$.

Theorem 5. *If H is a subgroup of a group G, then G is equal to the union of all right cosets of H in G i.e.,*

$G=H \cup Ha \cup Hb \cup Hc$......, where $a, b, c,$......are elements of G.

Proof. G is a group. Therefore each element of any right coset of H in G is an element of G. Hence the union of all right cosets of H in G is a subset of G.

Also if x is any element of G, then $x \in Hx$. Therefore x belongs to the union of all right cosets of H in G. Hence G is a subset of the union of all right cosets of H in G.

Therefore G is equal to the union of all right cosets of H in G. Symbolically, we have $G= \bigcup_{x \in G} Hx$.

Similarly we can prove that G is also equal to the union of all left cosets of H in G.

Right coset Decomposition of a group. Suppose H is a subgroup of a group G. No right coset of H in G is empty. Any two right cosets of H in G are either disjoint or identical. The union of all right cosets of H in G is equal to G. Therefore the set of all right cosets of H in G gives us a partition of G.

This partition is called the right coset decomposition of G with respect to the subgroup H. To obtain distinct members of this partition we should proceed as follows :

First of all H itself is a right coset. If there is an element $a \in G$ such that $a \notin H$, then Ha will be another distinct right coset. Again if there is an element $b \in G$ such that $b \notin H$ and also $b \notin Ha$, then Hb will be another distinct right coset. Proceeding in this way we can get all distinct right cosets of H in G. Then we shall have $G = H \cup Ha \cup Hb \cup Hc \ldots$, where $a, b, c, \ldots$ are elements of G so chosen that all right cosets are distinct.

Similarly we can also obtain left coset decomposition of G.

Theorem 6. *If H is a subgroup of G, there is a one-to-one correspondence between any two right cosets of H in G.*

Proof. Let $a, b \in G$. Then Ha and Hb are any two right cosets of H in G. Let $f : Ha \to Hb$ be defined by

$$f(ha) = hb \ \forall \ h \in H.$$

The function f is one-one. If $h_1, h_2 \in H$, then $h_1 a, h_2 a \in Ha$. Also by def. of f, we have $f(h_1 a) = h_1 b$ and $f(h_2 a) = h_2 b$.

Now
$$\begin{aligned} & f(h_1 a) = f(h_2 a) \\ \Rightarrow \ & h_1 b = h_2 b \\ \Rightarrow \ & h_1 = h_2 \quad \text{[by right cancellation law in } G] \\ \Rightarrow \ & h_1 a = h_2 a. \end{aligned}$$

$\therefore$ f is one-one since only equal elements of Ha can have the same image in Hb.

The function f is onto. Let $h'b$ be any arbitrary element of Hb. Then $h'b \in Hb \Rightarrow h' \in H \Rightarrow h'a \in Ha$. Now $f(h'a) = h'b$, by definition of f. Thus $h'b \in Hb \Rightarrow$ that there exists $h'a \in Ha$ such that $f(h'a) = h'b$. Therefore f is onto Hb.

Hence the result.

Similarly it can be proved that there is a 1-1 *correspondence between any two left cosets of H in G.* **(Kerala 1970)**

Note. H itself is a right as well as a left coset. Therefore if H is a finite subgroup of G, then *the number of elements in H i.e., $o(H)$ is equal to the number of elements in any coset of H in G (right or left).*

If H is an infinite subgroup of G, then we say that any two cosets of H in G have the same Cardinal number.

Theorem 7. *If H is a subgroup of G, then there is a one-to-one correspondence between the set of left cosets of H in G and the set of right cosets of H in G.* **(Kumayun 1977 ; Kanpur 80 ; Meerut 80)**

Proof. Let us define a function f from the set of left cosets of H in G to the set of right cosets of H in G by the formula

$$f(aH)=Ha^{-1} \;\forall\; a \in G.$$

First we shall show that this function f is well-defined. If aH is a left coset, then obviously Ha^{-1} is a right coset. Further if aH and bH represent the same left coset, then we are to show that $f(aH)=f(bH)$.

We have $aH=bH \Rightarrow a^{-1}b \in H$ [see theorem 2, page 154]

$$\Rightarrow Ha^{-1}b=H \qquad [\because h \in H \Rightarrow Hh=H]$$

$$\Rightarrow Ha^{-1}bb^{-1}=Hb^{-1} \Rightarrow Ha^{-1}=Hb^{-1}$$

$$\Rightarrow f(aH)=f(bH). \qquad [\because f(aH)=Ha^{-1} \text{ etc.}]$$

$\therefore$ f is well-defined.

Now to show that f is one-one. We have

$$f(aH)=f(bH)$$

$$\Rightarrow Ha^{-1}=Hb^{-1} \qquad \text{[by def. of } f]$$

$$\Rightarrow a^{-1}(b^{-1})^{-1} \in H \qquad \text{[See theorem 2, page 154]}$$

$$\Rightarrow a^{-1}b \in H$$

$$\Rightarrow a^{-1}bH=H \qquad [\because h \in H \Rightarrow hH=H]$$

$$\Rightarrow aa^{-1}bH=aH \Rightarrow bH=aH \Rightarrow aH=bH.$$

Hence f is one-one.

Now to show that f is onto the set of right cosets. Let Ha be any right coset. Then $a^{-1}H$ is a left coset. Also

$$f(a^{-1}H)=H(a^{-1})^{-1} \qquad \text{[by def. of } f]$$

$$=Ha.$$

Thus each right coset Ha is the f-image of the left coset $a^{-1}H$. Hence f is onto.

Note. From this theorem we conclude that if the number of distinct right cosets of H in G is finite, then it will also be equal to the number of distinct left cosets of H in G. It should be noted that in an infinite group G, it is possible that the number of distinct right cosets is finite.

Index of a subgroup in a group. Definition. *If H is a subgroup of a group G, the number of distinct right (left) cosets of H in G is called the index of H in G and is denoted by $[G : H]$ or by $i_G(H)$.*

[Punjab 1966 ; Banaras 63 ; Meerut 76, 83 (P)]

§ 28. Relation of congruence modulo a subgroup H in a group G.

Supppose H is a subgroup of a group G. If the element a of G belongs to the right coset Hb *i.e.*, if $a \in Hb$ *i.e.*, if $ab^{-1} \in H$, then we say that *a is congruent to b modulo H.*

Definition. *Let H be a subgroup of a group G. For a, b* $\in G$ *we say that a is congruent to b mod H if and only if* $ab^{-1} \in H$. *Symbolically, we write*

$$a \equiv b \ (\text{mod } H) \text{ if and only if } ab^{-1} \in H.$$

Theorem. *The relation of congruency in a group G defined by*

$$a \equiv b \ (\text{mod } H) \text{ iff } ab^{-1} \in H$$

is an equivalence relation.

(I.A.S. 1970; Meerut 68; Agra 70; Lucknow 69; Utkal 70)

Proof. Reflexivity. Let a be any element of G. Then $aa^{-1}=e \in H$ since H is a subgroup of G. Therefore $a \equiv a$ (mod H) for all $a \in G$. Hence the relation is reflexive.

Symmetry. We have $a \equiv b$ (mod H) $\Rightarrow ab^{-1} \in H$

$\Rightarrow (ab^{-1})^{-1} \in H$ [$\because$ H is a subgroup of G]

$\Rightarrow ba^{-1} \in H \Rightarrow b \equiv a$ (mod H).

Therefore the relation is symmetric.

Transitivity. Let $a \equiv b$ (mod H) and $b \equiv c$ (mod H). Then $ab^{-1} \in H$ and $bc^{-1} \in H$. But H is a subgroup of G and thus H must be closed with respect to the composition in G. Therefore

$$(ab^{-1})(bc^{-1}) \in H \Rightarrow a(b^{-1}b)c^{-1} \in H \Rightarrow aec^{-1} \in H$$

$$\Rightarrow ac^{-1} \in H \Rightarrow a \equiv c \ (\text{mod } H).$$

Hence the relation is transitive.

Therefore the relation congruence mod H is an equivalence relation in G. Therefore it will partition G into disjoint equivalence classes. If $\bar{a}$ or $[a]$ is the equivalence class corresponding to $a \in G$, then we shall show that $[a]=Ha$.

By the definition of equivalence classes, we have

$$[a]=\{x \in G : x \equiv a \ (\text{mod } H)\}.$$

Let z be any arbitrary element of Ha.

Then $z \in Ha \Rightarrow za^{-1} \in Haa^{-1} \Rightarrow za^{-1} \in He \Rightarrow za^{-1} \in H$

$\Rightarrow z \equiv a$ (mod H) $\Rightarrow z \in [a]$.

$\therefore$ $Ha \subseteq [a]$.

Now let y be any arbitrary element of $[a]$.

Then $y \in [a] \Rightarrow y \equiv a$ (mod H)

$\Rightarrow ya^{-1} \in H$

$\Rightarrow ya^{-1}a \in Ha \Rightarrow y \in Ha$.

$\therefore$ $[a] \subseteq Ha$.

Hence $[a]=Ha$.

Therefore the partition of G induced by this equivalence relation is nothing but **the right coset decomposition** of G with respect to H. No right coset of H in G will be empty. Two right cosets of H in G will be either disjoint or identical. The union of all right cosets of H in G will be equal to G.

Note. *We can show that the relation in G defined by*

$$a \equiv b \pmod{H} \text{ iff } a^{-1} b \in H$$

is an equivalence relation and gives us the left coset decomposition of G with respect to H.

Illustrations. 1. If G is the additive group of integers and H is the subgroup of G obtained on multiplying the elements of G by 5, then we have

$$G = H \cup (H+1) \cup (H+2) \cup (H+3) \cup (H+4).$$

The index of H in G is 5.

2. Let G be the group of all permutations of degree 3 on three symbols 1, 2, 3 and H be the subgroup $\{(1), (1\ 2)\}$. The number of elements in each right coset of H in G will be 2 and the number of elements in G is 6. So we shall have three distinct right cosets. It can be seen that $G = H \cup H(23) \cup H(31)$.

The left coset decomposition of G with respect to H is

$$G = H \cup (23)\ H \cup (31) H.$$

§ 29. Lagrange's theorem.

(I.A.S. 1973 ; Nagpur 70; Punjab 70; Kumayun 78; Poona 73; Kanpur 80; Banaras 71; Rajasthan 77; Nagarjuna 79; Luck. 80; Delhi 70; Madurai 78; Madras 83; Meerut 77, 79, 81, 82, 83, 84; Garhwal 76)

The order of each subgroup of a finite group is a divisor of the order of the group.

Proof. Let G be a group of finite order n. Let H be a subgroup of G and let $o(H) = m$ Suppose $h_1, h_2, \ldots, h_m$ are the m members of H.

Let $a \in G$. Then Ha is a right coset of H in G and we have

$$Ha = \{h_1 a, h_2 a, \ldots, h_m a\}.$$

Ha has m distinct members, since $h_i a = h_j a \Rightarrow h_i = h_j$.

Therefore each right coset of H in G has m distinct members. Any two distinct right cosets of H in G are disjoint *i.e.*, they have no element in common. Since G is a finite group, the number of distinct right cosets of H in G will be finite, say, equal to k. The union of these k distinct right cosets of H in G is equal to G. Thus if $\qquad Ha_1, Ha_2, \ldots, Ha_k$

are the k distinct right cosets of H in G, then

$$G=Ha_1 \cup Ha_2 \cup \ldots \cup Ha_k$$

$\Rightarrow$ the number of elements in $G=$ the number of elements in Ha_1+ the number of elements in $Ha_2+\ldots+$the number of elements in Ha_k. [$\because$ two distinct right cosets are mutually disjoint]

$$\Rightarrow o\ (G)=km \Rightarrow n=km$$

$$\Rightarrow k=\frac{n}{m} \Rightarrow m \text{ is a divisor of } n$$

$$\Rightarrow o\ (H) \text{ is a divisor of } o\ (G).$$

Hence the theorem.

Note 1. k is the index of H in G. We have $m=n/k$. Thus k is a divisor of n.

Therefore the index of every subgroup of a finite group is a divisor of the order of the group.

Note 2. Very Important. If H is a subgroup of a finite group G, then the index of H in $G=$the number of distinct right (or left) cosets of H in $G=\dfrac{o\ (G)}{o\ (H)}$.

Cor. 1. *The order of every element of a finite group is a divisor of the order of the group.* **(G N.D U. Amritsar 1982; Punjab 70; Poona 70; Banaras 70; Meerut 74, 76)**

Proof. Suppose G is a finite group of order n Let $a \in G$ and let $o\ (a)=m$. To prove that m is a divisor of n.

Let $H=\{\ldots, a^{-3}, a^{-2}, a^{-1}, a^0, a^1, a^2, a^3, \ldots\}$ be the subset of G consisting of all integral powers of a. Then we know that H is a subgroup of G. We shall show that H contains only m distinct elements and that they are $a, a^2, a^3, \ldots, a^m=e=a^0$.

Let $1 \leqslant r \leqslant m$, $1 \leqslant s \leqslant m$ and $r>s$.

Then $a^r=a^s \Rightarrow a^r a^{-s}=a^s a^{-s} \Rightarrow a^{r-s}=a^0 \Rightarrow a^{r-s}=e$.

Thus there exists a positive integer $r-s$ less than m such that $a^{r-s}=e$. But m is the least positive integer such that $a^m=e$. Therefore $a^r \neq a^s$. Therefore $a, a^2, a^3, \ldots, a^m=a^0=e$ are all distinct elements of H.

Now suppose a^t is any element of H, where t is any integer. By division algorithm, we have

$t=mp+q$ where p and q are some integers and $0 \leqslant q<m$.

[Note. We can write $t/m=p+q/m$].

We have $a^t=a^{mp+q}=a^{mp}a^q=(a^m)^p a^q=ea^q=a^q$. Since $0 \leqslant q<m$, therefore a^q is one of the m elements $a, a^2, \ldots, a^m=a^0$. Hence H has only m distinct elements.

Thus order of H is m. By Lagrange's theorem m is a divisor of n.

Cor. 2. *If G is a finite group of order n and $a \in G$, then $a^n = e$.*
(Madras 1970, Meerut 70, Punjab 70)

Proof. In a finite group, the order of each element is finite. Let $o(a) = m$. The subset H of G consisting of all integral powers of a is a subgroup of G and the order of H is m. By Lagrange's theorem m is a divisor of n. Let $k = \frac{n}{m}$. Then $n = mk$.

Now $$a^n = a^{mk} = (a^m)^k = e^k \quad [\because \; o(a) = m \Rightarrow a^m = e]$$ $$= e.$$

Note. Lagrange's theorem has very important applications. Suppose G is a finite group of order n. If m is not a divisor of n, then there can be no subgroup of G of order m. Thus if G is a group of order 6, then there can be no subgroup of G of order 5 or 4. Similarly if G is a group of prime order p, then G can have no proper subgroups. For suppose H is a proper subgroup of G and $o(H) = m$. Then $1 < m < p$ By Lagrange's theorem, m must be a divisor of p. But if $1 < m < p$ and p is prime, then m cannot divide p. So a group of prime order can have no proper subgroups.

However the converse of Lagrange's theorem is not true.
(Kanpur 1969)

If m is a divisor of n, then it is not necessary that G must have a subgroup of order m. For example the alternating group A_4 of degree 4 is of order 12. It can be seen that there is no subgroup of A_4 of order 6, though 6 is a divisor of 12.

Cor 3. Euler's Theorem. *If n is a positive integer and a is any integer relatively prime to n, then*

$$a^{\phi(n)} \equiv 1 \pmod{n},$$

where ϕ is the Euler ϕ-function.

Proof. For any integer x let $[x]$ denote the residue class of the set of integers mod n. Let $G = \{[a] : a \text{ is an integer relatively prime to } n\}$.

Then we know that with respect to multiplication of residue classes G is a group of order $\phi(n)$. The identity element of this group is the residue class $[1]$. We have

$$[a] \in G \Rightarrow [a]^{o(G)} = [1] \Rightarrow [a]^{\phi(n)} = [1]$$

$\Rightarrow [a]\,[a]\,[a] \ldots$ upto $\phi(n)$ times $=[1]$

$\Rightarrow [aa \ldots$ upto $\phi(n)$ times$]=[1]$ [Note that $[a]\,[b]=[ab]$]

$\Rightarrow [a^{\phi(n)}]=[1] \Rightarrow a^{\phi(n)} \equiv 1 \pmod{n}$.

Cor. 4. Fermat's Theorem. *If p is a prime number and a is any integer, then* $a^p \equiv a \pmod{p}$. **(Allahabad 1980)**

Proof. Let G be the set of non-zero residue classes of integers modulo p. If p is a prime number, then with respect to multiplication of residue classes G is a group of order $p-1$. The identity element of this group is $[1]$.

Now suppose a is any integer.

Case 1. p is a divisor of a. In this case $[a]=[0]$ and so $[a]$ is not an element of G. But

$$p \mid a \Rightarrow p \mid a^p \Rightarrow p \mid (a^p - a) \Rightarrow a^p \equiv a \pmod{p}.$$

Case 2. p is not a divisor of a. In this case $[a] \neq [0]$ and so $[a]$ is an element of G. Therefore we have

$[a]^{o(G)}=[1] \Rightarrow [a]^{p-1}=[1] \Rightarrow [a^{p-1}]=[1] \Rightarrow a^{p-1} \equiv 1 \pmod{p}$

$\Rightarrow a^{p-1}-1$ is divisible by $p \Rightarrow a\,(a^{p-1}-1)$ is divisible by p

$\Rightarrow a^p - a$ is divisible by $p \Rightarrow a^p \equiv a \pmod{p}$.

§ 30. Order of the product of two subgroups of finite order.

Theorem. *Let H and K be finite subgroups of a group G.*

Then $$o(HK)=\frac{o(H)\,o(K)}{o(H \cap K)}.$$

(I.A.S. 1973 ; Meerut 78, 82 P, 83 ; Andhra 75; Kanpur 80)

Proof. HK is a subset of G. It is not necessary that it will be a subgroup of G. By $o(HK)$ we mean the number of distinct elements in HK.

Let $D=H \cap K$. Then D is a subgroup of G and $D \subseteq K$. Therefore D is a subgroup of K. Since K is finite, therefore the number of distinct right cosets in the right coset decomposition of K with respect to D is finite. Let it be m. By Lagrange's theorem, we have $m=\dfrac{o(K)}{o(D)}$.

If $Dk_1, Dk_2, \ldots, Dk_m$ are the distinct right cosets of D in K,

then $$K=Dk_1 \cup Dk_2 \cup \ldots \cup Dk_m = \bigcup_{i=1}^{m} Dk_i.$$

Observe that $k_1, k_2, \ldots, k_m$ are some distinct elements in K.

Now $HK = H\left(\bigcup_{i=1}^{m} Dk_i\right) = \bigcup_{i=1}^{m} HDk_i$

$$= \bigcup_{i=1}^{m} Hk_i \qquad [\because\ D \subseteq H \Rightarrow HD = H]$$

$$= Hk_1 \cup Hk_2 \cup \ldots \cup Hk_m. \qquad \ldots(1)$$

We shall show that the cosets $Hk_1, Hk_2, \ldots, Hk_m$ are pairwise distinct. We have

$$Hk_i = Hk_j \Rightarrow k_i k_j^{-1} \in H$$
$$\Rightarrow k_i k_j^{-1} \in H \cap K \qquad [\because\ k_i, k_j \in K \Rightarrow k_i k_j^{-1} \in K]$$
$$\Rightarrow k_i k_j^{-1} \in D \Rightarrow Dk_i = Dk_j$$
$$\Rightarrow k_i = k_j \qquad [\because\ Dk_1, \ldots, Dk_m \text{ are distinct cosets}]$$

Thus $Hk_1, Hk_2, \ldots, Hk_m$ are distinct right cosets and so they are pairwise disjoint also. The number of elements in each of them is equal to $o(H)$ *i.e.*, the number of elements in H. Therefore from (1) we conclude that the number of elements in HK is equal to $m \times o(H)$.

$$\therefore\ o(HK) = m \times o(H) = \frac{o(K)}{o(D)} \cdot o(H) = \frac{o(H)\, o(K)}{o(H \cap K)}.$$

Corollary. *Let H and K be subgroups of a finite group G and let $o(H) > \sqrt{[o(G)]}$, $o(K) > \sqrt{[o(G)]}$. Then $H \cap K \neq \{e\}$.*

(Meerut 1980)

Proof. Since $HK \subseteq G$, therefore $o(HK) \leqslant o(G)$. $\ldots(1)$

But $$o(HK) = \frac{o(H)\, o(K)}{o(H \cap K)}. \qquad \ldots(2)$$

From (1) and (2), we get $o(G) \geqslant \dfrac{o(H)\, o(K)}{o(H \cap K)}$. $\ldots(3)$

But $\dfrac{o(H)\, o(K)}{o(H \cap K)} > \dfrac{\sqrt{[o(G)]}\sqrt{[o(G)]}}{o(H \cap K)}$ [by hypothesis]

i.e., $$\frac{o(H)\, o(K)}{o(H \cap K)} > \frac{o(G)}{o(H \cap K)} \qquad \ldots(4)$$

From (3) and (4), we get $o(G) > \dfrac{o(G)}{o(H \cap K)}$.

Therefore $o(H \cap K) > 1$ and this implies that $H \cap K \neq \{e\}$ because order of $\{e\} = 1$.

We apply this corollary to a very special group.

Example. *Let G be a finite group of order pq where p and q are prime numbers with $p > q$. Then G has at most one subgroup of order p. In particular a group of order 6 has at most one subgroup of order 3.* **(Meerut 1972, 76)**

Solution. If possible, let H and K be the two subgroups of G of the same order p. Since $p > q$ and $o(G)=pq$, therefore

$$o(H) > \sqrt{[o(G)]} \text{ and } o(K) > \sqrt{[o(G)]}.$$

So by the above corollary $H \cap K \neq \{e\}$.

Now $H \cap K$ is a subgroup of H. Since H is of prime order p, therefore either $H \cap K=H$ or $H \cap K=\{e\}$. But $H \cap K \neq \{e\}$. Therefore $H \cap K=H$ and this implies that $H \subseteq K$. Similarly we can prove that $K \subseteq H$. Hence $H=K$. Therefore there can be at most one subgroup of G of order p.

As a special case a group of order six has at most one subgroup of order 3 since $6=3 \times 2$ where $3 > 2$ and both 3 and 2 are primes.

Solved Examples

Ex. 1. *Show that two right cosets Ha, Hb are distinct if and only if the two left cosets $a^{-1}H$, $b^{-1}H$ are distinct.*

Solution. In order to prove the given statement we shall prove that two right cosets Ha, Hb are equal if and only if the two left cosets $a^{-1}H$, $b^{-1}H$ are equal. We have

$$\begin{aligned} Ha=Hb &\Leftrightarrow ab^{-1} \in H && \text{[See theorem 2, page 154]} \\ &\Leftrightarrow ab^{-1}H=H && [\because h \in H \Leftrightarrow hH=H] \\ &\Leftrightarrow a^{-1}ab^{-1}H=a^{-1}H \\ &\Leftrightarrow b^{-1}H=a^{-1}H \Leftrightarrow a^{-1}H=b^{-1}H. \end{aligned}$$

Hence the required result follows.

Ex. 2. *Show that the set of the inverses of the elements of a right coset is a left coset; or more precisely show that $(Ha)^{-1}=a^{-1}H$.*

Solution. Suppose Ha is a right coset of H in G where $a \in G$.

Let ha be any element of Ha, where $h \in H$.

We have $(ha)^{-1}=a^{-1}h^{-1}$.

Since H is a subgroup, therefore $h \in H \Rightarrow h^{-1} \in H$.

$$\therefore \quad a^{-1}h^{-1} \in a^{-1}H.$$

Thus the inverses of all the elements of Ha belong to the left coset $a^{-1}H$. Hence $(Ha)^{-1} \subseteq a^{-1}H$.

Conversely, let $a^{-1}h$ be any element of $a^{-1}H$.

Then $a^{-1}h=a^{-1}(h^{-1})^{-1}=(h^{-1}a)^{-1} \in (Ha)^{-1}$,
since $h^{-1} \in H$ and therefore $h^{-1}a \in Ha$.

Therefore every element of $a^{-1}H$ belongs to the set of the inverses of the elements of Ha.

$$\therefore \quad a^{-1}H \subseteq (Ha)^{-1}.$$

Hence $(Ha)^{-1}=a^{-1}H$.

Ex. 3. *Given that* $G=H\cup Ha_2\cup Ha_3\cup\ldots\cup Ha_k$ *is the right coset decomposition of* G *relative to the subgroup* H*, show that* $G=H\cup a_2^{-1}H\cup a_3^{-1}H\cup\ldots\cup a_k^{-1}H$ *is a left coset decomposition of* G *relative to the subgroup* H.

Solution. We know that two right cosets of H in G are either disjoint or identical. Therefore if there are two equal right cosets in the given right coset decomposition of G, then one of them can be omitted. So let us assume that all the right cosets in the given right coset decomposition of G relative to H are distinct. Then there are k distinct right cosets of H in G. But the number of distinct right cosets of H in G is equal to the number of distinct left cosets of H in G. Therefore there are k distinct left cosets in the left coset decomposition of G relative to H.

Now we know that two right cosets Ha and Hb are distinct if and only if the two left cosets $a^{-1}H$ and $b^{-1}H$ are distinct. Since the right cosets $H, Ha_2, \ldots, Ha_k$ are all distinct, therefore the left cosets $H, a_2^{-1}H, \ldots, a_k^{-1}H$ are all distinct. Since they are k in number, therefore they are the only distinct left cosets of H in G. Now the union of all the distinct left cosets of H in G is equal to G. Hence

$$G=H\cup a_2^{-1}H\cup a_3^{-1}H\cup\ldots\cup a_k^{-1}H$$

is a left coset decomposition of G relative to the subgroup H.

Ex. 4 *Prove that the only right (or left) coset of a subgroup* H *in a group* G *which is also a subgroup of* G *is* H *itself.*

Solution. Suppose Ha is a right coset of H in G. Let Ha be a subgroup of G. Then $e\in Ha$. But $e\in H$. Since H is itself a right coset and two right cosets are either disjoint or identical, therefore $H=Ha$.

Similarly aH is a subgroup of $G\Rightarrow aH=H$.

Ex. 5. *If* $H\subseteq K$ *are two subgroups of a finite group* G*, then show that* $[G:H]=[G:K]\,[K:H]$.

Solution. Since $H\subseteq K$ are two subgroups of a group G, therefore H is also a subgroup of K.

Now H is a subgroup of a finite group G. Therefore by Lagrange's theorem

$$[G:H]=\frac{o(G)}{o(H)}=\frac{o(G)}{o(K)}\cdot\frac{o(K)}{o(H)}=[G:K]\,[K:H].$$

Ex. 6. *Let* H *and* K *be two subgroups of a group* G*. Show that any coset relative to* $H\cap K$ *is the intersection of a coset relative to* H *with a coset relative to* K. **(Jabalpur 1970)**

Solution. Let a be any element of G. Then $(H\cap K)\,a$ is any right coset of G relative to the subgroup $H\cap K$. We shall prove that

$$(H\cap K)\,a=(Ha)\cap(Ka).$$

We have $(H\cap K)\subseteq H \Rightarrow (H\cap K)\,a\subseteq Ha$

and $(H\cap K)\subseteq K \Rightarrow (H\cap K)\,a\subseteq Ka.$

$\therefore\quad (H\cap K)\,a\subseteq Ha\cap Ka.$ ...(1)

Again let x be any element of $Ha\cap Ka$. Then $x\in Ha$ and $x\in Ka$.

$\therefore\quad x=ha=ka$ for some $h\in H$, $k\in K$.

$xa^{-1}=h=k$.

$\therefore\quad xa^{-1}\in H\cap K \Rightarrow (xa^{-1})\,a\in(H\cap K)\,a \Rightarrow x\in(H\cap K)\,a.$

Consequently $Ha\cap Ka\subseteq(H\cap K)\,a$. ...(2)

From (1) and (2), we conclude that $(H\cap K)\,a=Ha\cap Ka$.

A similar proof can be given in the case of a left coset.

Ex. 7. *Prove that the intersection of two subgroups, each of finite index, is again of finite index.* **(Meerut 1973 ; Madras 83)**

Solution. Let H and K be two subgroups of a group G. Let $[G:H]=m$ and $[G:K]=n$. Let $Ha_1,\ldots,Ha_m$ and $Kb_1,\ldots,Kb_n$ be the distinct right cosets of H and K respectively. We are to show that the number of distinct right cosets of $H\cap K$ in G is finite. Let $(H\cap K)\,a$ be any right coset of $H\cap K$ in G. Then it can be easily shown that $(H\cap K)\,a=Ha\cap Ka$. Thus each right coset of $H\cap K$ is given by the intersection of a right coset of H and a right coset of K. Since the number of distinct right cosets of H is m and the number of distinct right cosets of K is n, therefore the number of distinct right cosets of $H\cap K$ can be at most equal to mn. Hence $H\cap K$ is of finite index in G.

Ex. 8. *Use Lagrange's theorem to prove that a finite group cannot be expressed as the union of two of its proper subgroups.* **(Poona 1973)**

Solution. Let G be a finite group of order n. Suppose G is the union of two of its proper subgroups H and K.

Since $e\in$ both H and K and $G=H\cup K$, therefore at least one of H and K (say, H) must contain more than half the elements of G. Let $o(H)=p$. Then $n/2<p<n$. (Note that H is a proper subgroup of G).

Since $n/2<p<n$, therefore p cannot be a divisor of n. This contradicts Lagrange's theorem which states that the order of each subgroup of a finite group is a divisor of the order of the group.

Hence our initial assumption is wrong and so a finite group cannot be expressed as the union of two of its proper subgroups.

§ 31. Cayley's Theorem. *Every finite group G is isomorphic to a permutation group.* **(G.N.D.U. Amritsar 1982 ; I.A.S. 69 ; Allahabad 68 ; Delhi Hons. 69 ; Punjab 69 ; Kanpur 80 ; Madras 83 ; Meerut 78, 81, 83 P)**

Proof. Let G be a finite group. If $a \in G$, then for every x in G the product ax is also an element of G. Now consider the function f_a from G into G defined by

$$f_a(x) = ax \ \forall \ x \in G.$$

The function f_a is one-one because if $x, y \in G$, then

$$f_a(x) = f_a(y) \Rightarrow ax = ay$$
$$\Rightarrow x = y. \quad \text{[by left cancellation law in } G\text{]}$$

The function f_a is also onto because if x is any element of G, then $\exists$ an element $a^{-1}x$ in G such that

$$f_a(a^{-1}x) = a(a^{-1}x) = (aa^{-1})x = ex = x.$$

Thus f_a is a one-one function from G onto G. Therefore f_a is a permutation on G. Let G' denote the set of all such one-one onto functions defined on G corresponding to every element of G *i.e.*,

$$G' = \{f_a : a \in G\}.$$

First we shall show that G' is a group with respect to the operation known as composite or product of two functions.

Closure property. Let $f_a, f_b \in G'$ where $a, b \in G$. From our definition of product of two functions, we have

$$(f_a f_b)(x) = f_a[f_b(x)] = f_a(bx) = a(bx) = (ab)x$$
$$= f_{ab}(x) \text{ for all } x \in G.$$

Therefore by the definition of equality of two functions, we have

$$f_a f_b = f_{ab}. \qquad \ldots(1)$$

Since $ab \in G$, therefore $f_{ab} \in G'$ and thus G' is closed with respect to the product of functions.

Associativity. Let $f_a, f_b, f_c \in G'$ where $a, b, c \in G$. Then

$$\begin{aligned} f_a(f_b f_c) &= f_a f_{bc} && [\because \text{ from (1)}, f_b f_c = f_{bc}] \\ &= f_{a(bc)} && [\text{from (1)}] \\ &= f_{(ab)c} && [\text{by associativity in } G] \\ &= f_{ab} f_c && [\text{from (1)}] \\ &= (f_a f_b) f_c. && [\text{from (1)}] \end{aligned}$$

Therefore the operation in G' is associative.

Existence of Identity. If e is the identity of G, then f_e is the identity of G' because for every f_a in G', we have

$$f_e f_a = f_{ea} = f_a \text{ and } f_a f_e = f_{ae} = f_a.$$

Existence of Inverse. If a^{-1} is the inverse of a in G, then $f_{a^{-1}}$ is the inverse of f_a in G' because

$$f_{a^{-1}} f_a = f_{a^{-1}a} = f_e \text{ and } f_a f_{a^{-1}} = f_{aa^{-1}} = f_e.$$

Thus G' is a group.

Now we shall show that $G \cong G'$. Consider the function ϕ from G into G' defined by $\phi(a) = f_a \ \forall \ a \in G$.

ϕ is one-one. If $a, b \in G$, then

$$\begin{aligned} \phi(a) = \phi(b) &\Rightarrow f_a = f_b \Rightarrow f_a(x) = f_b(x) \ \forall \ x \in G \\ &\Rightarrow ax = bx \ \forall \ x \in G \Rightarrow a = b. \end{aligned}$$

$\therefore$ ϕ is one-one.

ϕ is onto. Let f_a be any element of G'.

Then $a \in G$ and we have $\phi(a) = f_a$. Therefore ϕ is onto.

ϕ preserves compositions in G and G'. If $a, b \in G$, then

$$\begin{aligned} \phi(ab) &= f_{ab} && \text{[by def. of } \phi] \\ &= f_a f_b && \text{[from (1)]} \\ &= \phi(a)\,\phi(b). && \text{[by def. of } \phi] \end{aligned}$$

$\therefore$ ϕ preserves compositions in G and G'.

$\therefore$ $G \cong G'$.

Definition. *The permutation group to which G is isomorphic is called a regular permutation group.*

Note 1. While forming the function f_a we have multiplied the elements of G on the left hand side by a and not on the right hand side. Thus we have taken $f_a(x) = ax$ and not $f_a(x) = xa$. If we take $f_a(x) = xa$, we shall get $f_a f_b = f_{ba}$. Then while defining the mapping $\phi : G \to G'$ we should set $\phi(a) = f_{a^{-1}}$ and not $\phi(a) = f_a$.

Note 2. Important. Cayley's theorem is true even if the group G is not finite. The same proof can be given in that case because while proving the theorem we have not assumed that G is finite.

But if G is not finite, then the word permutation should be omitted from the statement of the theorem. In that case we should state the theorem like this :

Every group is isomorphic to a group of one-one onto functions.

(Sagar 1966, 68 ; Dibrugarh 67 ; Jabalpur 69)

Example. *Find the regular permutation group isomorphic to the multiplicative group* $G = \{1, -1, i, -i\}$. **(Raj. 1978)**

Solution. By Cayley's theorem the regular permutation group G' isomorphic to G consists of the following four permutations f_1, f_{-1}, f_i, f_{-i} :

$$f_1=\begin{pmatrix}1 & -1 & i & -i\\ 1\cdot 1 & (1)\cdot(-1) & (1)\cdot(i) & (1)\cdot(-i)\end{pmatrix}=\begin{pmatrix}1 & -1 & i & -i\\ 1 & -1 & i & -i\end{pmatrix}=I$$

$$f_2=\begin{pmatrix}1 & -1 & i & -i\\ (-1)\cdot 1 & (-1)(-1) & (-1)\,i & (-1)(-i)\end{pmatrix}$$

$$=\begin{pmatrix}1 & -1 & i & -i\\ -1 & 1 & -i & i\end{pmatrix}=(1,\ -1)\,(i,\ -i).$$

$$f_3=\begin{pmatrix}1 & -1 & i & -i\\ (i)\,1 & (i)\,(-1) & i\cdot i & (i)\,(-i)\end{pmatrix}$$

$$=\begin{pmatrix}1 & -1 & i & -i\\ i & -i & -1 & 1\end{pmatrix}=(1,\ i,\ -1,\ -i).$$

$$f_4=\begin{pmatrix}1 & -1 & i & -i\\ (-i)\,1 & (-i)\,(-1) & (-i)\,i & (-i)\,(-i)\end{pmatrix}$$

$$=\begin{pmatrix}1 & -1 & i & -i\\ -i & i & 1 & -1\end{pmatrix}=(1,\ -i,\ -1,\ i).$$

Exercises

1. What is a coset ? Give an example. **(Meerut 1972)**

2. What are the left cosets and right cosets of a subgroup H of a group G ? Prove that any two left cosets of H in G are identical or have no elements in common. Show that there is a 1-1 correspondence between any two left cosets of H in G.

(Kerala 1970)

3. Correct the following statement :

If G is a group and H is a subgroup of G, then $o\,(H)$ is a divisor of $o\,(G)$. **(Meerut 1976)**

4. Answer the following questions :

(*i*) What is the order of the group P_4 ?

(*ii*) What is the order of the group A_4 ?

(*iii*) Is A_4 a subgroup of P_4 ? If yes, what is the index of A_4 in P_4 ? **Ans.** (i) 24, (ii) 12, (iii) 2.

5. Use Fermat's theorem to determine the remainder, if 8^{103} is divided by 103. **(Meerut 1977)**

Ans. 8.

6. Prove that any group is isomorphic to a transformation group. **(Kumayun 1977 ; Allahabad 80)**

7. (i) Show that every group is isomorphic to a subgroup of $A(S)$ for some appropriate S. Here $A(S)$ is the group of all one-one functions of S onto itself. **(Meerut 1973 ; Vikram 76)**

(ii) Let G be a finite group. Prove that G is isomorphic to a subgroup of S_n, the symmetric group of degree n, for some n.

(G.N.D.U. Amritsar 1982)

8. Let G be the group of integers under addition and let N be the set of all integral multiples of 3. Prove that N is a sub-

group of G and determine all the cosets of N in G. (I.A.S. 1973)

9. Let G be a finite group, $a \in G$; show that the order of a equals the order of the subgroup H of G generated by a. Hence or otherwise deduce that $o(a)$ divides $o(G)$. **(Meerut 1975)**

10. Consider two subgroups $H = \{I, (1\ 2)\}$ and $K = \{I, (1\ 3)\}$ of S_3. Determine HK. Using Lagrange's theorem or otherwise prove that HK is not a subgroup of S_3. **(Meerut 1976)**

11. If H and K are subgroups of a finite group G, give an example to show that $o(HK)$ need not divide $o(G)$. **(Meerut 1973)**

12 If a finite group G contains an element of even order, show that G must also be of even order.

13. If a finite group possesses an element of order 2, prove that it possesses an odd number of such elements.

14. S is a subset of a given group G and one defines a relation $a \sim b$ in G if and only if $ab^{-1} \in S$. Show that the necessary and sufficient condition that this is an equivalence relation is that S is a subgroup of G. (I.A.S. 1975)

15. Use Lagrange's theorem to show that any group of prime order can have no proper subgroups.

§ 32. Cyclic groups.

Definition. *A group G is called cyclic if, for some $a \in G$, every element $x \in G$ is of the form a^n, where n is some integer. The element a is then called a generator of G.*

(I.A.S. 1974; Meerut 81; Garhwal 76; Madras 74; Kumayon 78; Madras 77; Bombay 70; Rajasthan 77)

There may be more than one generators of a cyclic group. If G is a cyclic group generated by a, then we shall write $G = \{a\}$ or $G = (a)$. The elements of G will be of the form

$$\ldots, a^{-3}, a^{-2}, a^{-1}, a^0 = e, a, a^2, a^3, \ldots$$

Of course they are not necessarily all distinct.

Example 1. **(Vikram 1978).** *The multiplicative group $G = \{1, -1, i, -i\}$ is cyclic. We can write $G = \{i, i^2, i^3, i^4\}$. Thus G is a cyclic group and i is a generator. Also we can write*

$$G = \{-i, (-i)^2, (-i)^3, (-i)^4\}.$$

Thus $-i$ is also a generator of G.

Example 2. The multiplicative group $\{1, \omega, \omega^2\}$ is cyclic. The generators are ω and ω^2.

Example 3. Suppose G is any group and $a \in G$. Let H be the subgroup of G consisting of all integral powers of a *i.e.*,

$H=\{a^n : n \in I$ (the set of integers)$\}$. Then H is a cyclic subgroup of G generated by a.

Example 4. The group $G=(\{0, 1, 2, 3, 4, 5\}, +_6)$ is cyclic. This group is generated by 1. Another generator is 5.

We see that $1^1=1$, $1^2=1+_6 1=2$, $1^3=1+_6 1^2=3$,

$$1^4=1+_6 1^3=1+_6 3=4,\ 1^5=1+_6 1^4=5,\ 1^6=0.$$

Thus $\quad G=\{1, 1^2, 1^3, 1^4, 1^5, 1^6=0\}$.

Example 5. The multiplicative group of n nth roots of unity is cyclic, a generator being $e^{2\pi i/n}$. **(Allahabad 70; Raj. 78)**

Example 6. The additive group G of all integers is cyclic, a generator being 1. We have $1^0=0$, $1^1=1$, $1^2=1+1=2$, $1^3=1+1+1=3$ and so on. Similarly $1^{-1}=$ inverse of $1=-1$, $1^{-2}=(1^2)^{-1}=-2$, $1^{-3}=(1^3)^{-1}=(3)^{-1}=-3$ and so on.

Thus each element of G can be expressed as some integral power of 1. Also -1 is a generator.

§ 33. Some Properties of Cyclic Groups.

Theorem 1. *Every cyclic group is an abelian group.*

(Meerut 1976; Allahabad 80; Vikram 78)

Proof. Let $G=\{a\}$ be a cyclic group generated by a. Let x, y be any two elements of G. Then there exist integers r and s such that $x=a^r$, $y=a^s$. Now $xy=a^r\, a^s=a^{r+s}=a^{s+r}=a^s\, a^r=yx$. Thus we have $xy=yx \ \forall\ x, y \in G$. Therefore G is abelian.

Theorem 2. *If a is a generator of a cyclic group G, then a^{-1} is also a generator of G.*

Proof. Let $G=\{a\}$ be a cyclic group generated by a. Let a^r be any element of G, where r is some integer. We can write $a^r=(a^{-1})^{-r}$. Since $-r$ is also some integer, therefore each element of G, is generated by a^{-1}. Thus a^{-1} is also a generator of G.

Theorem 3. *A cyclic group G with generator of finite order n, is isomorphic to the multiplicative group of n nth roots of unity.*

(Meerut 1971; Poona 70; Madurai 78)

Proof. Let a be a generator of the cyclic group G. Since the order of a is n, therefore n is the least positive integer such that

$$a^n=e.$$

We shall show that the group G has exactly n distinct elements

$$a,\ a^2,\ a^3,\ \ldots,\ a^n=e=a^0. \qquad \ldots(1)$$

No two elements of (1) can be equal. For if possible, let

$$a^r=a^s,\ 1 \leqslant s < r \leqslant n. \quad \text{Then } a^{r-s}=a^0=e.$$

Since $0<r-s<n$, therefore $a^{r-s}=e$ implies that the order of a is less than n. Hence $a^r \neq a^s$.

Therefore all the n elements in (1) are distinct.

Again let a^t be any element of G. By division algorithm, there exist two integers p and q such that

$t=np+q,\ 0 \leqslant q < n.$ [Note. We can write $t/n=p+q/n$]

$$\therefore\ a^t=a^{np+q}=a^{np}a^q=(a^n)^p a^q=e^p a^q=ea^q=a^q.$$

Since $0 \leqslant q < n$, therefore a^q is one of the n elements in (1). Thus each element of G is equal to some member of (1). Therefore G has exactly n elements given in (1). Thus $o(a)=o(G)$.

We shall now show that G is isomorphic to the multiplicative group G' of the n nth roots of unity, namely

$$1=e^{2\pi i0/n},\ e^{2\pi i/n},\ e^{4\pi i/n},\ \ldots,\ e^{2\pi ri/n},\ \ldots,\ e^{2\pi(n-1)i/n}.$$

Consider the mapping $f: G \rightarrow G'$ defined by

$$f(a^r)=e^{2\pi ir/n}, \text{ where } 0 \leqslant r \leqslant n-1.$$

The mapping f is one-one. Since

$f(a^r)=f(a^s)$, where $0 \leqslant r \leqslant n-1,\ 0 \leqslant s \leqslant n-1$

$$\Rightarrow e^{2\pi ir/n}=e^{2\pi is/n} \Rightarrow r=s \Rightarrow a^r=a^s.$$

Again the number of elements in G is equal to the number of elements in G'. Therefore f is one-one implies f must be onto G'.

Finally $f(a^r\, a^s)=f(a^{r+s})=f(a^{nu+k})$, where u is some integer and $0 \leqslant k < n$

[Note. We can write $(r+s)/n=u+k/n$]

$=f(a^{nu}a^k)=f[(a^n)^u\, a^k]$

$=f(a^k)$ [$\because$ $a^n=e$]

$=e^{2\pi ki/n}$ [by def. of f]

$=e^{2\pi nui/n}\, e^{2\pi ki/n}$ [$\because$ $e^{2\pi ui}=1$]

$=e^{2\pi i(nu+k)/n}=e^{2\pi i(r+s)/n}=e^{2\pi ir/n}\, e^{2\pi is/n}=f(a^r)\, f(a^s).$

Therefore f preserves compositions in G and G'. Hence G is isomorphic to G'.

Since every finite cyclic group of order n is isomorphic to the multiplicative group of n nth roots of unity, therefore we can say that *there is one and only one cyclic group of order n*.

Cor. If a is a generator of an infinite cyclic group G, then the order of a must be infinite. If the order of a is finite, then the cyclic group generated by a will be of finite order. Therefore *the order of a cyclic group is equal to the order of its generating element.* **(Rohilkhand 1980)**

Theorem 4. *A cyclic group G with a generator of finite order n, is isomorphic to the additive group of residue classes modulo n.*
(Punjab 1970; Jiwaji 78; Meerut 73)

Proof. First to prove that the group G has exactly n distinct elements, give the same proof as in theorem 3. The group G' is here the group of residue classes modulo n.

For any integer a let $[a]$ denote the residue class of the set of integers modulo n. Then $G'=\{[a] : a \in \text{I}$ where I is the set of integers$\}$. The group G' has only n distinct elements and we have $G'=\{[0], [1], [2], \ldots, [n-1]\}$. Also $G=\{a^r : r \in \text{I}\}$.

Consider the mapping $f : G \to G'$ defined by

$$f(a^r)=[r] \ \forall \ r \in \text{I}.$$

First we must show that the mapping f is **well-defined.**

Let $r, s \in \text{I}$ be such that $a^r=a^s$.

Then we must show that $f(a^r)=f(a^s)$.

We have $a^r=a^s \Rightarrow a^r a^{-s}=a^s a^{-s} \Rightarrow a^{r-s}=e$

$$\Rightarrow n \text{ is a divisor of } r-s \qquad [\because o(a)=n]$$
$$\Rightarrow r \equiv s \ (\text{mod } n) \Rightarrow [r]=[s] \Rightarrow f(a^r)=f(a^s).$$

$\therefore$ the mapping f is well-defined.

f is **one-one.** Let a^r, a^s be any two elements of G where $r, s \in \text{I}$.

We have $f(a^r)=f(a^s) \Rightarrow [r]=[s] \Rightarrow r-s$ is divisible by n

$\Rightarrow r-s=kn$ where $k \in \text{I}$

$\Rightarrow a^{r-s}=a^{kn} \Rightarrow a^r a^{-s}=(a^n)^k$

$\Rightarrow a^r a^{-s}=e^k \Rightarrow a^r a^{-s}=e \Rightarrow a^r=a^s \Rightarrow f$ is one-one.

f is **onto.** Let $[r]$ be any element of G'. Then r is an integer. We have $a^r \in G$ and $f(a^r)=[r]$. Therefore f is onto G'.

f **preserves compositions.**

We have $f(a^r a^s)=f(a^{r+s})=[r+s]=[r]+[s]=f(a^r)+f(a^s)$.

$\therefore \ G \cong G'$.

Theorem 5. *If a finite group of order n contains an element of order n, the group must be cyclic.* **(Banaras 1970)**

Proof. Suppose G is a finite group of order n. Let $a \in G$ and let n be the order of a. If H is the cyclic subgroup of G generated by a *i.e.*, if $H=\{a^r : r \in \text{I}\}$, then the order of H is n because the order of the generator a of H is n. Thus H is a cyclic subgroup of G and the order of H is equal to the order of G. Hence $H=G$ and therefore G itself is a cyclic group and a is a generator of G.

Note. Suppose G is a finite group of order n and we are to find whether G is cyclic or not. We should find the orders of the elements of G. If we are able to find an element $a \in G$ such that $o(a)=n$, then G will be a cyclic group and a will be a generator of G.

Theorem 6. *Every group of prime order is cyclic.*

(I.A.S. 1972; Kanpur 80; Guru Nanak 75; Madurai 78; Rajasthan 76; Meerut 76, 80, 81; Gorakhpur 70)

Proof. Suppose G is a finite group whose order is a prime number p, then to prove that G is a cyclic group. Note that an integer p is said to be a prime number if $p \neq 0$, $p \neq \pm 1$, and if the only divisors of p are ± 1, $\pm p$.

Since G is a group of prime order, therefore G must contain at least 2 elements. Note that 2 is the least positive prime integer. Therefore there must exist an element $a \in G$ such that $a \neq$ the identity element e.

Since a is not the identity element, therefore $o(a)$ is definitely >2. Let $o(a)=m$. Then $H=\{a\}$ is a cyclic subgroup of G and $o(H)=o(a)=m$. By Lagrange's theorem m must be a divisor of p. But p is prime and $m \geqslant 2$. Hence $m=p$.

$\therefore$ $H=G$. Since H is cyclic therefore G is cyclic and a is a generator of G.

Note. Every group of prime order p is cyclic. Also every cyclic group of order p is isomorphic to the additive group of residue classes modulo p. Therefore we can say that if p is prime then there is one and only one group of order p.

Theorem 7. *Every finite group of composite order possesses proper subgroups.* **(Delhi Hons. 1970; Allahabad 80; Meerut 70)**

Proof. Let G be a finite group of composite order mn where neither m nor n is 1. Note that a non-zero integer α is said to be a composite number if there exist two integers β, γ such that $\alpha=\beta\gamma$ while $|\beta| > 1$, $|\gamma| > 1$.

Suppose G is cyclic and a is a generator of G. Then

$$o(a)=o(G)=mn.$$

$\therefore$ $a^{mn}=e$

$$\Rightarrow (a^n)^m=e \Rightarrow o(a^n) \text{ is finite and is } \leqslant m.$$

Let $o(a^n)=r$ where $r<m$. Then $(a^n)^r=e \Rightarrow a^{nr}=e$.
But $r<m \Rightarrow nr<mn$.

Thus $a^{nr}=e$ where $nr<mn$ while $o(a)=mn$. This is not possible because mn is the least positive integer such that $a^{mn}=e$.

$\therefore$ We must have $r=m$. Thus $o(a^n)=m$.

$\therefore$ $H=\{a^n\}$ is a cyclic subgroup of G and the order of H is equal to the order of its generator a^n. Thus $o(H)=m$. Since $2\leqslant m<mn$, therefore H is a proper subgroup of G.

Suppose G is not cyclic. Then the order of each element of G must be less than mn. So G will contain at least one element, say b, of order p where $2\leqslant p<mn$. Then $H=\{b\}$ is a subgroup of G and $o(H)=p$. Therefore H is a proper subgroup of G.

Theorem 8. *If a cyclic group G is generated by an element a of order n, then a^m is a generator of G if and only if the greatest common divisor of m and n is* 1 *i.e., iff m and n are relative primes.*

(Bombay 1970; Calicut 75; Kumayon 78; Meerut 74, 79; Nagarjuna 79; Rajasthan 77)

Proof Suppose m is relatively prime to n. Consider the cyclic subgroup $H=\{a^m\}$ of G generated by a^m. Obviously $H\subseteq G$ since each integral power of a^m will also be an integral power of a.

Since m is relatively prime to n, therefore there exist two integers u, v such that $um+vn=1$.

$$\therefore\ a^{um+vn}=a^1$$
$$\Rightarrow a^{um}\,a^{vn}=a$$
$$\Rightarrow (a^m)^u=a;\text{ since } a^{vn}=(a^n)^v=e^v=e.$$

$\therefore$ each integral power of a will also be some integral power of a^m. Therefore $G\subseteq H$. Hence $H=G$ and a^m is a generator of G.

Converse. Suppose a^m is a generator of G. Let the greatest common divisor of m and n be d and $d\neq 1$ i.e., $d>1$. Then m/d and n/d must be integers.

Now $(a^m)^{n/d}=(a^n)^{m/d}=e^{m/d}=e$. Obviously n/d is a positive integer less than n itself. Thus $o(a^m)<n$. Therefore a^m cannot be a generator of G because the order of a^m is not equal to the order of G. Hence d must be equal to 1. Thus m is prime to n.

Note. If G is a cyclic group of order n, then the total number of generators of G will be equal to the number of integers less than n and prime to n. For example if a is a generator of a cyclic group G of order 8, then a^3, a^5, a^7 will be the only other generators of G. Since 4 is not prime to 8, therefore a^4 cannot be a generator of G. Similarly a^2, a^6, a^8 cannot be generators of G.

If G is a cyclic group of prime order p generated by a, then $a, a^2, \ldots, a^{p-1}$ are all generators of G and thus G has $p-1$ generators.

Theorem 9. *If G is an infinite cyclic group, then G has exactly two generators and G is isomorphic to the additive group of integers.*

(I.A.S. 1974; Punjab 76; Sagar 77; Madras 74; Meerut 75)

Proof. Let $G=\{a\}$ be an infinite cyclic group generated by a. The elements of G will be integral powers of a. We claim that no two distinct integral powers of a can be equal.

For, if possible, let $a^r=a^s$, $r>s$.

Then $$a^r a^{-s}=a^s a^{-s}=a^0=e.$$

$$\therefore \quad a^{r-s}=e.$$

Since $r-s$ is a positive integer, therefore $a^{r-s}=e$ implies that $o(a)$ is finite. So a cannot be a generator of an infinite cyclic group G. Hence $a^r \neq a^s$, unless $r=s$.

Therefore we can write $G=\{\ldots, a^{-3}, a^{-2}, a^{-1}, a^0=e, a, a^2, a^3, \ldots\}$.

If a^r is any element of G, we can write $a^r=(a^{-1})^{-r}$. Thus a^{-1} is also a generator of G. Also as proved above a and a^{-1} are distinct elements of G.

Now if $m \neq 1$ or -1, then a^m cannot be a generator of G. If a^m is to be a generator of G, there must exist an integer k such that $(a^m)^k=a$, *i.e.*, $a^{mk}=a$. Now $m \neq 1$ or -1, $\Rightarrow mk \neq 1$. Therefore two distinct integral powers of a are equal and this contradicts the statement we have just proved. Hence a^m cannot be a generator of G if $m \neq 1$ or -1. Thus G has exactly two generators *i.e.*, a and a^{-1}. Let I be the additive group of integers *i.e.*,

$$\mathrm{I}=\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}.$$

To prove that $\quad G \cong \mathrm{I}$.

Consider the mapping $\phi : G \to \mathrm{I}$ defined by

$$\phi(a^m)=m \ \forall \ m \in \mathrm{I}.$$

The mapping ϕ is one-one because

$$\phi(a^m)=\phi(a^n) \Rightarrow m=n \Rightarrow a^m=a^n.$$

Obviously ϕ is onto I.

Also $\phi(a^m a^n)=\phi(a^{m+n})=m+n=\phi(a^m)+\phi(a^n)$.

Hence ϕ is an isomorphism of G onto I.

Note. Since every infinite cyclic group is isomorphic to the additive group of integers, therefore we can say that there is one and only one infinite cyclic group.

Theorem 10. *Every subgroup of a cyclic group is cyclic.*

(Punjab 1970; Osmania 72; Allahabad 79, Bombay 70; Indore 70; Madurai 75; Madras 74; Meerut 73, 78)

Proof Suppose $G=\{a\}$ is a cyclic group generated by a. If $H=G$ or $\{e\}$, then obviously H is cyclic. So let H be a proper subgroup of G. The elements of H are integral powers of a. If $a^s \in H$, then the inverse of a^s *i.e.*, $a^{-s} \in H$. Therefore H contains elements which are positive as well as negative integral powers of a. Let m be the least positive integer such that $a^m \in H$. Then we shall prove that $H=\{a^m\}$ *i.e.*, H is cyclic and is generated by a^m.

Let a^t be any arbitrary element of H. By division algorithm, there exist integers q and r such that $t=mq+r$, $0 \leqslant r < m$.

Now $\quad a^m \in H \Rightarrow (a^m)^q \in H \quad$ [by closure property]

$$\Rightarrow a^{mq} \in H \Rightarrow (a^{mq})^{-1} \in H$$
$$\Rightarrow a^{-mq} \in H.$$

Also $a^t \in H,\ a^{-mq} \in H \Rightarrow a^t a^{-mq} \in H \Rightarrow a^{t-mq} \in H$

$$\Rightarrow a^r \in H. \qquad [\because\ r=t-mq]$$

Now m is the least positive integer such that $a^m \in H$ and $0 \leqslant r < m$. Therefore r must be equal to 0. Hence $t=mq$.

$$\therefore\quad a^t=a^{mq}=(a^m)^q.$$

Thus every element $a^t \in H$ is of the form $(a^m)^q$. Therefore H is cyclic and a^m is a generator of H.

Theorem 11. *Every proper subgroup of an infinite cyclic group is infinite.*

Proof. Let $G=\{a\}$ be an infinite cyclic group. Let H be a proper subgroup of G. Then H is cyclic and if m is the least positive integer such that $a^m \in H$, then $H=\{a^m\}$.

Suppose H is a finite group of order p. Since a^m is a generator of H, therefore $(a^m)^p=e$ *i e.*, $a^{mp}=e$ where $mp > 0$.

$\therefore$ Order of a is finite and consequently G is finite, contrary to the hypothesis. Hence H must be an infinite cyclic subgroup of G.

Note. Since every infinite cyclic group is isomorphic to the group of integers, therefore any two proper subgroups of an infinite cyclic group are isomorphic, even isomorphic to the group itself.

Solved Examples

Ex. 1. *Show that the group* $(\{1, 2, 3, 4, 5, 6\}, \times_7)$ *is cyclic. How many generators are there ?*

Solution. Let us denote the given group by G. If there exists an element $a \in G$ such that $o(a)=6$ *i.e.*, equal to the order of the group G then the group G will be a cyclic group and a will be a generator of G.

We see that $o(3)=6$ because $3^1=3$, $3^2=3 \times_7 3=2$, $3^3=3^2 \times_7 3=2 \times_7 3=6$, $3^4=6 \times_7 3=4$, $3^5=4 \times_7 3=5$, $3^6=5 \times_7 3=1$ *i.e.*, the identity element.

$\therefore$ G is cyclic and 3 is a generator of G. We can write

$$G=\{3, 3^2, 3^3, 3^4, 3^5, 3^6\}.$$

Now 5 is prime to 6. Therefore 3^5 *i.e.*, 5 is also a generator of G.

Ex. 2. *How many generators are there of the cyclic group G of order 8 ?*

Solution. Let a be a generator of G. Then $o(a)=8$. We can write $G=\{a, a^2, a^3, a^4, a^5, a^6, a^7, a^8\}$.

7 is prime to 8, therefore a^7 is also a generator of G.

5 is prime to 8, therefore a^5 is also a generator of G.

3 is prime to 8, therefore a^3 is also a generator of G.

Thus there are only four generators of G *i.e.*, a, a^3, a^5, a^7.

Ex. 3. *Show that every isomorphic image of a cyclic group is again cyclic.* **(Punjab 1970; Vikram 78)**

Solution. Let $G=\{a\}$ be a cyclic group generated by a. Let G' be an isomorphic image of G under the isomorphism f. The elements of G' are the images of the elements of G under the function f. Let $f(a^n)$ be any element of G'. We have

$f(a^n)=f(aaa \ldots \text{to } n \text{ factors})=f(a)\, f(a)\, f(a) \ldots \text{to } n \text{ factors}=[f(a)]^n$.

Thus we see that every element of G' can be expressed as an integral power of $f(a)$. Thus G' is cyclic and $f(a)$ is a generator of G'.

Ex. 4. *If $G=\{a\}$ be a finite cyclic group of order n, then for any divisor d of n, there exists a unique subgroup of G of order d.*

(Banaras 1963; Meerut 76)

Solution. Existence. Since d is a divisor of n, therefore we have $n=dm$ for some integer m.

Since $G=\{a\}$ is of order n, therefore

$$o(a)=n$$

$$\Rightarrow a^n=e$$

$$\Rightarrow a^{dm}=e \qquad [\because n=dm]$$

$$\Rightarrow (a^m)^d=e \Rightarrow o(a^m) \leqslant d.$$

Let $o(a^m)=s$ where $s<d$. Then

$(a^m)^s=e \Rightarrow a^{ms}=e$ where $ms<md$ i.e., $ms<n$.

Now $o(a)=n$. Therefore if $ms<n$, then we cannot have $a^{ms}=e$.

$\therefore$ $o(a^m)$ cannot be less than d and it must be equal to d.

Now $a^m \in G$ and $o(a^m)=d$. Therefore $\{a^m\}$ is a cyclic subgroup of G of order d.

Uniqueness. We know that every subgroup of a cyclic group is cyclic.

If possible suppose there is another subgroup $\{a^k\}$ of G of order d where $n=dm$.

We shall show that $\{a^k\}=\{a^m\}$. By division algorithm, there exist integers q and r such that

$$k=mq+r \text{ where } 0 \leqslant r<m. \qquad \ldots(1)$$

Therefore $kd=mqd+rd$ where $0 \leqslant rd<md$.

So $$a^{kd}=a^{mqd+rd}=a^{mqd}a^{rd}=(a^{md})^q\, a^{rd}=(a^n)^q\, a^{rd}$$
$$=e^q a^{rd}=a^{rd}. \qquad \ldots(2)$$

Since the subgroup $\{a^k\}$ is of order d, therefore its generator a^k is also of order d. Therefore $a^{kd}=e$.

So from (2), we get

$$a^{rd}=e,$$

which is impossible unless $r=0$ because $rd<md$ i.e., $rd<n$ and $o(a)=n$. Putting $r=0$ in (1), we get

$$k=mq.$$

$$\therefore \quad a^k=a^{mq}=(a^m)^q.$$

$$\therefore \quad a^k \in \{a^m\}.$$

$$\therefore \quad \{a^k\} \subseteq \{a^m\}. \qquad \ldots(3)$$

But number of elements of $\{a^k\}$ = number of elements of $\{a^m\}$.

$\therefore$ (3) implies $\{a^k\}=\{a^m\}$.

Ex. 5. *Let G be a group, $G \neq \{e\}$. Then G has no proper subgroups if, and only if, G is a finite cyclic group of prime order.*

(Punjab 1970)

Solution. 'If' part. Let G be a finite group of prime order p. Then G is cyclic. By Lagrange's theorem the order of every subgroup H of G must be a divisor of the order of G. So if H is any subgroup of G, we must have either $o(H)=1$ or $o(H)=p$. Note that p is prime. Thus G can have no proper subgroups.

'Only if' part. Suppose G has no proper subgroups. Then to prove that G must be a finite cyclic group of prime order. Let $x \in G$ be such that $x \neq e$. Then the cyclic subgroup $\{x\}$ of G generated by x is all of G because it is given that G has no proper subgroups. Note that $x \neq e \Rightarrow \{x\} \neq \{e\}$. Thus G is a cyclic group generated by x. Now if x is of infinite order, then x^2 generates a proper subgroup of the cyclic group G and this again contradicts the hypothesis that G has no proper subgroups. So $o(x)$ must be finite, say n. Since $x \neq e$, therefore $n > 1$. Suppose n is not prime. Then we must have $n = rs$ where r and s are some positive integers such that $1 < r < n$, $1 < s < n$. Now $o(x) = n$ and r is a positive integer which divides n. So we must have $o(x^r) = \frac{n}{r} = s$. Then x^r generates a proper subgroup of G of order s which again contradicts our hypothesis. So n must be prime. Thus G is a cyclic group whose generator x is of prime order n. But the order of a cyclic group is equal to the order of its generator. Hence G must be a finite cyclic group of prime order.

Ex. 6. *Show that every finite group of order less than six must be abelian.* **(Meerut 1971)**

Solution. We know that every group of prime order is cyclic and every cyclic group is abelian. Since 2, 3 and 5 are prime numbers, therefore all groups of order 2, 3 and 5 must be abelian.

To show that every group of order 4 must be abelian, see example 12 on page 120.

Ex. 7. *Taking a group $\{e, a, b, c\}$ of order 4, construct two composition tables which are not isomorphic.* **(Meerut 1980, 82)**

Solution. Every group of order 4 is abelian. But a group of order 4 may be cyclic or it may not be cyclic Suppose $\{e, a, b, c\}$ is a group of order 4. It must contain an element, other than e, which is its own inverse. Let $a^{-1} = a$. Then $a^2 = e$ and $o(a) = 2$.

Now two cases arise. If $b^{-1} = b$, we must have $c^{-1} = c$ In this case the group contains no element of order 4 and so it is not cyclic.

But if $b^{-1} = c$, then we must have $c^{-1} = b$. Also in this case $o(b)$ cannot be 2 and so it must be 4. Thus in this case the group is cyclic.

The composition tables for the two groups of order four are as given below :

Cyclic group of order 4

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
	b	c	a	e
	c	b	e	a

Non-cyclic group of order 4

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

The two types of groups given above are not isomorphic. We see that one group is cyclic while the other is not cyclic,

Ex. 8. *Give an example of a finite abelian group which is not cyclic.* **(Meerut 1979, 83P)**

Solution. Let G be the set of the four real matrices

$$I=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix},\ A=\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix},\ B=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix},\ C=\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}.$$

It can be easily seen that G is an abelian group with respect to multiplication of matrices. The identity element of this group is the identity matrix I. Let us find the order of each element of G. We have $o(I)=1$. Also

$$A^2=\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=I;\quad \therefore\ o(A)=2;$$

$$B^2=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=I;\quad \therefore\ o(B)=2;$$

$$C^2=\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=I;\quad \therefore\ o(C)=2.$$

Now G is a group of order 4 and G contains no element of order 4. Therefore G is not a cyclic group. Hence G is a finite abelian group which is not cyclic.

Ex. 9. *Let a and b be two elements of finite order, of a group G. If $o(a)$ and $o(b)$ are co-prime and $ab=ba$, prove that*

$$o(ab)=o(a)\,o(b).$$

Solution. Suppose $o(a)=m$ and $o(b)=n$, where m and n are relatively prime. Consider the cyclic subgroup H of G generated by ab. We have

$$\begin{aligned}(ab)^{mn}&=a^{mn}b^{mn} && [\because\ ab=ba]\\ &=(a^m)^n(b^n)^m=e^ne^m && [\because\ o(a)=m\Rightarrow a^m=e \text{ and similarly } b^n=e]\\ &=e.\end{aligned}$$

$\therefore \quad o(ab) \mid mn.$

But $o(ab)=o(H)$, ab being a generator of the cyclic group H.

$$\therefore \quad o(H) \mid mn. \qquad \ldots(1)$$

Again ab is a generator of the cyclic group $H \Rightarrow (ab)^m \in H \Rightarrow a^m b^m \in H \Rightarrow eb^m \in H \Rightarrow b^m \in H$. Since b is of order n and m is relatively prime to n, therefore $o(b^m)=n$. But we have just shown that $b^m \in H$. Therefore by a corollary to Lagrange's theorem, $o(b^m) \mid o(H)$ *i.e.*, $n \mid o(H)$. Similarly we can show that $m \mid o(H)$. But m and n are relatively prime. Therefore $m \mid o(H)$ and $n \mid o(H)$

$$\Rightarrow mn \mid o(H). \qquad \ldots(2)$$

From (1) and (2), we conclude that $o(H)=mn$. Consequently $o(ab)=mn$.

Ex. 10. *If an abelian group of order six contains an element of order* 3, *show that it must be a cyclic group.*

Solution. Let G be an abelian group of order 6. Let $a \in G$ be such that $o(a)=3$. To prove that G is cyclic. Since G is of order six (even), therefore G must contain an element b other than e such that $b^{-1}=b$ *i.e.*, $b^2=e$ *i.e.*, $o(b)=2$. We shall show that $o(ab)=6$ and consequently G will be a cyclic group generated by ab.

We have $b^{-1} \neq a$, since $o(b^{-1})=o(b)=2$ while $o(a)=3$. Thus $ab \neq e$. Now $(ab)^2=a^2b^2=a^2e=a^2 \neq e$, since $o(a)=3$. Again $(ab)^3 =a^3b^3=eb^2b=eeb=b \neq e$. Therefore we must have $o(ab) > 3$. But $o(ab)$ must be a divisor of $o(G)$ *i.e.*, 6. So $o(ab)$ can neither be 4 nor it can be 5. Hence we must have $o(ab)=6$ and consequently G is cyclic.

Exercises

1. Show that any two cyclic groups of the same order are isomorphic. **(Jabalpur 1970 ; Marathwada 70 ; Mysore 70)**

2. Show that all finite cyclic groups of order n are isomorphic to the additive group of integers modulo n. **(Kerala 1970)**

3. Find the order of each element in the multiplicative group of residues 1, 2, 3, 4, 5, 6 prime to 7. Show that the group is cyclic of order 6 and that it can be generated by 3 and 5 and not by any of the other elements. **(Kerala 1970)**

4. "A group may be isomorphic to one of its proper subgroups". Either disprove this statement or give an example to prove it. **(Nagpur 1970)**

5. Explain the statement, "every element of a group generates a subgroup". **(Nagpur 1970)**

6. Define a cyclic group

Let a be a generator of a cyclic group G. Let $\phi : n \rightarrow a^n$ be a mapping of the additive group of all integers **I** onto G. (i) 'If ϕ is one-one, prove that G is isomorphic to **I**, (ii) If ϕ is not one-one, prove that G is a finite cyclic group. **(Kerala 1970)**

7. Give an example to show that a group of order 4 is not necessarily cyclic.

8. Give an example of an infinite non-abelian group and one example of an abelian group which is not cyclic. **(Mysore 1970)**

9. If ω be the cube root of unity, show that the set $\{1, \omega, \omega^2\}$ is a cyclic group of order 3 with respect to multiplication.

(Meerut 1975)

10. Show that the set U_n of n nth complex roots of unity forms a cyclic group with respect to multiplication. **(Allahabad 1970)**

11. Prove that a non-commutative group has at least six elements. **(Banaras 1971)**

12. What is the least order of a non-abelian group ? Show that all proper subgroups of a group of order 8 must be abelian.

(Meerut 1976, 80)

13. Show that the group $(\{1, 2, 3, 4\}, \times_5)$ is cyclic.

14. Show that the residue classes $\{1\}, \{3\}, \{5\}, \{7\}$ modulo 8 form a multiplicative group. Is this a cyclic group ? **Ans.** No.

(Kanpur 1970)

15. (i) How may generators are there of the cyclic group of order 10 ? **Ans.** Four *i.e.*, a, a^3, a^7, a^9.

(ii) How many generators can a cyclic group of order 12 have ? **Ans.** 4. **[Meerut 1980, 82 (P) ; Rajasthan 77]**

16. How many elements of the cyclic group of order 6 can be used as generators of the group ? **Ans.** Two *i.e.*, a and a^5.

17. Prove that each element $b \neq e$ in an infinite cyclic group $\{a\}$ is of infinite order.

18. Verify that the six matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix},$$

$$\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$$

form a group for multiplication of matrices. Show that this group is cyclic and that it may be generated by either of the last two matrices.

19. Define a cyclic group. Let $G=(a)$ be a finite cyclic group of order n. Show that the number of subgroups of G equals the number of divisors of n. **(Meerut 1976)**

20. Show that if G is a group of prime order p, then G is cyclic and has $(p-1)$ generators. **(I.A.S. 1972)**

21. Explain the notion of isomorphism of groups.

Prove or disprove the following statements :

(i) Any infinite cyclic group is isomorphic to the group **I** of integers under addition.

(ii) Any two finite groups with the same number of elements are isomorphic. **(I.A.S. 1972)**

22. Let G be a group having no proper subgroups. Show that G must be a finite group of order n where n is either 1 or a prime number. **(Punjab 1970)**

23. Prove that the order of a cyclic group is equal to the order of its generator.

24. Show that the number of generators of a finite cyclic group of order m is $\phi(m)$ where ϕ is the Euler's ϕ-function. **(Meerut 1979)**

25. Prove that a cyclic group with only one generator can have at most two elements. **(Meerut 1980)**

26. If the generator of a cyclic group G is of infinite order (or of order zero), prove that G is isomorphic to the additive group of integers.

§ 34. Subgroup generated by a subset of a group.

Definition. *Let M be any arbitrary subset of a group G. Then a subgroup H of G is called the smallest subgroup of G containing M if H contains M and if H is contained in every subgroup of G containing M.* **(Osmania 1972)**

The smallest subgroup of G containing M is called the **subgroup generated** by M and will be denoted by $\{M\}$ or by (M).

Theorem 1. *Let M be any subset of a group G. Prove that the smallest subgroup of G containing M exists and is unique.*

Proof. Let $\{M\}$ be the smallest subgroup of G containing M. Let F be the family of all subgroups of G which contain M. The family F is not empty because at least G belongs to this family. It should be noted that G itself is a subgroup of G and G definitely contains M. Let H be the intersection of the family F. Then H is a subgroup of G. Also obviously $M \subseteq H$.

Let K be any subgroup of G containing M. Then K is a member of the family F. Since H is the intersection of the family F, therefore $H \subseteq K$.

$\therefore$ $H=\{M\}$ *i.e.*, H is the smallest subgroup of G containing M.

Uniqueness. Let H_1, H_2 be two smallest subgroups of G containing M. Then we have $H_1 \subseteq H_2$ and $H_2 \subseteq H_1$.

$$\therefore \quad H_1=H_2.$$

Theorem 2. *Let M be a subset of a group G. Then the set of elements of G expressible as products of positive and negative integral powers of finite numbers of elements of M is the smallest subgroup of G containing M.*

Proof. The proof is easy. Let H be the set of those elements of G which can be expressed as products of positive and negative integral powers of finite numbers of elements of M. Let $a, b \in H$. Then obviously $ab^{-1} \in H$. Therefore, H is a subgroup of G. Obviously $M \subseteq H$. Also if K is any subgroup of G containing M, then definitely H must be contained in K Therefore,

$$H=\{M\}.$$

§ 35. Generating system of a group.

Definition. *A set M of a group G is said to be a generating system of G, if the subgroup generated by M coincides with G i.e.,*

$$G=\{M\}.$$

Every arbitrary group possesses at least one generating system, namely, the group itself.

Independent generating system. *A generating system is said to be* **irreducible** *or* **independent** *if no proper subset of it can generate G.* Thus if $G=\{M\}$ and if no member of M belongs to the subgroup of G generated by the set of remaining members of M, then M is an independent generating system of G.

It can be easily seen that every finite group possesses an independent generating system. An infinite group may or may not possess an independent generating system. *A finite group may possess several distinct independent generating systems containing different numbers of elements.* It will be clear from the following example.

Example. Let G be the cyclic group of order 6.

Let $G=\{a, a^2, a^3, a^4, a^5, a^6=e\}$.

The elements a and a^5 are both generators of G. We have

$$G=\{a\},\ G=\{a^5\}.$$

Both these are independent generating systems of G. Also

$$G=\{a^2, a^3\}.$$

The two elements a^2, a^3 also form an independent generating system because any of these elements does not lie in the subgroup generated by the other. The subgroup generated by a^2 is $\{a^2, a^4, a^6=e\}$ and a^3 is not in it. The subgroup generated by a^3 is $\{a^3, a^6=e\}$ and a^2 is not in it.

Solved Examples

Ex. 1. *Show that the set of permutations*

$$(1\ 2), (1\ 3), \ldots, (1\ n)$$

is an independent generating set for the symmetric group P_n of permutations of degree n.

Solution. We know that every permutation in P_n can be expressed as the product of disjoint cycles.

Now we shall show that every cycle can be expressed as product of members of the given set. Then the given set will generate P_n.

For a cycle containing the symbol 1, we have

$$(1\ a_1\ a_2 \ldots a_k)=(1\ a_1)\ (1\ a_2) \ldots (1\ a_k).$$

For a cycle not containing the symbol 1, we have

$$\begin{aligned}(b_1\ b_2 \ldots b_l) &= (1\ b_1)\ (1\ b_2\ b_3 \ldots b_l\ b_1)\\ &= (1\ b_1)\ (1\ b_2)\ (1\ b_3) .. (1\ b_l)\ (1\ b_1).\end{aligned}$$

Thus every permutation in P_n can be expressed as a product of members of the given set.

Also we see that each member of the given set of permutations contains an element not contained in any of the others. Therefore P_n cannot be generated by a proper subset of the given set of permutations. Hence the given set of permutations is an independent generating system of P_n.

Ex. 2. *Show that the $(n-1)$ permutations*

$$(1\ 2), (2\ 3), \ldots, (n-1\ n)$$

constitute a set of independent generators of P_n.

Solution. We have

$$(2\ 3)\ (1\ 2)\ (2\ 3)=(1\ 3),$$
$$(3\ 4)\ (1\ 3)\ (3\ 4)=(1\ 4),$$
$$\dots \quad \dots \quad \dots \quad \dots$$
$$\dots \quad \dots \quad \dots \quad \dots$$
$$(n-1\ n)\ (1\ n-1)\ (n-1\ n)=(1\ n).$$

Therefore each member of the set of permutations

$$(1\ 2),\ (1\ 3),\ ..,\ (1\ n)$$

is generated by the given set of permutations.

Now as shown in exercise 1, the set of permutations

$$(1\ 2),\ (1\ 3),\dots,\ (1\ n)$$

generates P_n.

Therefore the set of permutations

$$(1\ 2),\ (2\ 3),\dots,\ (n-1\ n)$$

also generates P_n.

Also we see that each member of the given set of permutations contains an element not contained in any of the others. Therefore P_n cannot be generated by a proper subset of the given set of permutations. Hence the given set of permutations is an independent generating system of P_n.

Ex. 3. *Let* S_3 *be the group of permutations of the set* $\{1, 2, 3\}$. *Prove that* $\{(1\ 2), (1\ 2\ 3)\}$ *generates* S_3. **(Meerut 1973)**

3

Groups (Continued)

§ 1. **Normal subgroups.** Let G be an abelian group the composition in G being denoted multiplicatively. Let H be any subgroup of G. If x is any element of G, then Hx is a right coset of H in G and xH is a left coset of H in G. Also G is abelian, therefore we must have $Hx=xH \ \forall \ x \in G$. However it is possible that G is not abelian, yet it possesses a subgroup H such that $Hx=xH \ \forall \ x \in G$. Such subgroups of a group G come under the category of **normal subgroups** and these are very important.

Normal subgroup. Definition.

(I. A. S. 1970; Guru Nanak 82; Nagarjuna 78, 80; Meerut 81, 82, 83, 84; Kanpur 80; Allahabad 80)

A subgroup H of a group G is said to be a normal subgroup of G if for every $x \in G$ and for every $h \in H$, $xhx^{-1} \in H$.

From this definition we can immediately conclude that H is a normal subgroup of G if and only if $xHx^{-1} \subseteq H \ \forall \ x \in G$.

Remark. We have $x \in G \Rightarrow x^{-1} \in G$. Therefore H is a normal subgroup of G iff $x^{-1} h (x^{-1})^{-1}$ *i.e.*, $x^{-1} hx \in H \ \forall \ x \in G$ and $\forall \ h \in H$.

Every group G possesses at least two normal subgroups namely G itself and the subgroup consisting of the identity element e alone. These are called improper normal subgroups. There exist groups for which these are the only normal subgroups. Such groups are known as **simple groups**.

Simple Group. Definition. *A group having no proper normal subgroups is called a simple group.*

Every group of prime order is simple. By Lagrange's theorem such a group has no proper subgroups.

Theorem 1. *A subgroup H of a group G is normal if and only if $xHx^{-1}=H \ \forall \ x \in G$.* **(Rajasthan 1978; Meerut 84)**

Proof. Let $xHx^{-1}=H \ \forall \ x \in G$. Then $xHx^{-1} \subseteq H \ \forall \ x \in G$. Therefore H is a normal subgroup of G.

Converse. Let H be a normal subgroup of G.

Then $\quad xHx^{-1} \subseteq H \ \forall \ x \in G$. ...(1)

Also $\quad x \in G \Rightarrow x^{-1} \in G$. Therefore we have

$x^{-1} H (x^{-1})^{-1} \subseteq H \ \forall \ x \in G \Rightarrow x^{-1} Hx \subseteq H \ \forall \ x \in G$

$\Rightarrow x\,(x^{-1}Hx)\,x^{-1} \subseteq xHx^{-1}\ \forall\ x \in G$

$\Rightarrow H \subseteq xHx^{-1}$ for all $x \in G$. ...(2)

From (1) and (2), we conclude that $xHx^{-1}=H$ for all $x \in G$.

Theorem 2. *A subgroup H of a group G is a normal subgroup of G if and only if each left coset of H in G is a right coset of H in G.* **(Kanpur 1980; I A.S. 72; Nagarjuna 79; Vikram 76; Meerut 77, 79, 82)**

Proof. Let H be a normal subgroup of G.

Then $xHx^{-1}=H\ \forall\ x \in G \Rightarrow (xHx^{-1})\,x=Hx$ for all $x \in G$

$\Rightarrow xH=Hx$ for all $x \in G$

$\Rightarrow$ each left coset xH is the right coset Hx.

Conversely suppose that each left coset of H in G is a right coset of H in G. Let x be any element of G. Then

$xH=Hy$, for some $y \in G$.

Since $e \in H$, therefore $xe=x \in xH$.

$\therefore\ x \in Hy$. $[\because\ Hy=Hx]$

But $x \in Hy \Rightarrow Hx=Hy$ [See theorem 3, page 154]

$\therefore\ Hx=xH$. $[\because\ Hy=xH]$

Thus we have

$xH=Hx\ \forall\ x \in G \Rightarrow xHx^{-1}=Hxx^{-1}\ \forall\ x \in G$

$\Rightarrow xHx^{-1}=H\ \forall\ x \in G \Rightarrow H$ is a normal subgroup of G.

Thus H is a normal subgroup of $G \Leftrightarrow xH=Hx\ \forall\ x \in G$.

Theorem 3. *A subgroup H of a group G is a normal subgroup of G if and only if the product of two right cosets of H in G is again a right coset of H in G.*

(Kanpur 1971; Andhra 77; Meerut 79, 80, 83)

Proof. Let H be a normal subgroup of a group G. Let a, b be any two elements of G. Then Ha and Hb are two right cosets of H in G. We have

$(Ha)\,(Hb)=H\,(aH)\,b$

$=H\,(Ha)\,b$ $[\because\ H$ is normal $\Rightarrow Ha=aH]$

$=HHab=Hab.$ $[\because\ HH=H]$

Since $a \in G, b \in G \Rightarrow ab \in G$, therefore Hab is also a right coset of H in G. Thus the product of the right cosets Ha and Hb is the right coset Hab.

Converse. Let H be a subgroup of G such that the product of two right cosets of H in G is again a right coset of H in G. Let x be any element of G. Then $x^{-1} \in G$. Therefore Hx and Hx^{-1} are two right cosets of H in G. Consequently, by hypothesis $HxHx^{-1}$ is also a right coset of H in G. Since $e \in H$, therefore $exex^{-1}=e$

is an element of the right coset $HxHx^{-1}$. But H itself is a right coset of H in G and $e\in H$. Also if two right cosets have one element common they must be identical. Therefore we must have

$$HxHx^{-1}=H \ \forall\ x\in G$$
$$\Rightarrow h_1xhx^{-1}\in H \ \forall\ x\in G \text{ and } \forall\ h_1, h\in H$$
$$\Rightarrow h_1^{-1}(h_1xhx^{-1})\in h_1^{-1}H \ \forall\ x\in G \text{ and } \forall\ h_1, h\in H$$
$$\Rightarrow xhx^{-1}\in H \ \forall\ x\in G \text{ and } \forall\ h\in H$$
$$[\because\ h_1^{-1}H=H \text{ as } h_1^{-1}\in H \text{ since } h_1\in H]$$
$$\Rightarrow H \text{ is a normal subgroup of } G.$$

Theorem 4. *The intersection of any two normal subgroups of a group is a normal subgroup.*

(Madras 1983; Rajasthan 71; Sambalpur 77)

Proof. Let H and K be two normal subgroups of a group G. Since H and K are subgroups of G, therefore $H\cap K$ is also a subgroup of G. Now to prove that $H\cap K$ is a normal subgroup of G. Let x be any element of G and n be any element of $H\cap K$. We have $n\in H\cap K\Rightarrow n\in H, n\in K$. Since H is a normal subgroup of G, therefore $x\in G, n\in H\Rightarrow xnx^{-1}\in H$. Similarly $xnx^{-1}\in K$.

Now $xnx^{-1}\in H, xnx^{-1}\in K\Rightarrow xnx^{-1}\in H\cap K$.

Thus we have $x\in G, n\in H\cap K\Rightarrow xnx^{-1}\in H\cap K$.

Hence $H\cap K$ is a normal subgroup of G.

Theorem 5. *The intersection of any collection of normal subgroups is itself a normal subgroup.*

(Raj. 1976; Meerut 80; Kanpur 80)

Proof. Let G be a group and let $\{H_t : t\in T\}$ be any family of normal subgroups of G. Here T is an index set and is such that $\forall\ t\in T$, H_t is a normal subgroup of G.

Let $H=\bigcap_{t\in T} H_t=\{a\in G : a\in H_t\ \forall\ t\in T\}$

be the intersection of this family of normal subgroups of G. Then to prove that H is also a normal subgroup of G.

We know that the intersection of any collection of subgroups is itself a subgroup. Therefore H is a subgroup of G. Now to prove that H is a normal subgroup of G.

Let x be any element of G and h be any element of $H=\bigcap_{t\in T} H_t$.

We have $h\in \bigcap_{t\in T} H_t\Rightarrow h\in H_t\ \forall\ t\in T$. Since H_t is a normal subgroup of G, therefore $x\in G, h\in H_t\Rightarrow xhx^{-1}\in H_t$.

Thus we have $x\in G, h\in \bigcap_{t\in T} H_t\Rightarrow xhx^{-1}\in H_t\ \forall\ t\in T$

$$\Rightarrow xhx^{-1} \in \bigcap_{t \in T} H_t \Rightarrow \bigcap_{t \in T} H_t \text{ is a normal subgroup of } G.$$

Solved Examples

Ex. 1. *Show that every subgroup of an abelian group is normal.*

(Kanpur 1980; Meerut 79; Guru Nanak 75)

Solution. Let G be an abelian group and H a subgroup of G. Let x be any element of G and h any element of H. We have

$$xhx^{-1} = xx^{-1}h \qquad [\because G \text{ is abelian} \Rightarrow x^{-1}h = hx^{-1}]$$
$$= eh = h \in H.$$

Thus $x \in G$, $h \in H \Rightarrow xhx^{-1} \in H$. Hence H is normal in G.

Note. Since every cyclic group is abelian, therefore every subgroup of a cylic group is normal.

Ex. 2 *Let P_n be the symmetric group on n symbols. Prove that A_n is a normal subgroup of P_n.* **(Rajasthan 1977; Nagarjuna 79)**

Solution. Let α be any element of P_n and β any element of A_n. Then β is an even permutation and α may be odd or even. We claim that $\alpha\beta\alpha^{-1}$ is an even permutation.

If α is odd, then α^{-1} is also odd. Now $\alpha\beta$ is odd and consequently $\alpha\beta\alpha^{-1}$ is even.

If α is even, then α^{-1} is also even. Now $\alpha\beta$ is even and consequently $\alpha\beta\alpha^{-1}$ is even.

Thus $\alpha \in P_n$, $\beta \in A_n \Rightarrow \alpha\beta\alpha^{-1} \in A_n$. Hence A_n is a normal subgroup of P_n.

Ex 3. *If G is a group and H is a subgroup of index 2 in G, prove that H is a normal subgroup of G.*

(Nagarjuna 1978; I.A S. 74; Punjab 70; Poona 73; Bombay 70; Kanpur 80; Raj. 78; Meerut 82P, 83P, 84P)

Solution. Let H be a subgroup of index 2 in a group G. Then the number of distinct right (left) cosets of H in G is 2. Let x be any element of G. If $x \in H$, then we have

$$xH = H = Hx.$$

If $x \notin H$, then the right coset Hx is distinct from H and the left coset xH is distinct from H. But H is of index 2; therefore the number of distinct right (left) cosets in right (left) coset decomposition of G will be 2. Therefore the cosets H, Hx, xH are such that $G = H \cup Hx = H \cup xH$. But there is no element common to H and Hx and also there is no element common to H and xH Therefore we must have $Hx = xH$. Thus we have $Hx = xH \; \forall \; x \in G$. Hence H is a normal subgroup of G.

Ex. 4. *H is a normal subgroup of G and K is a subgroup of G such that $H \subseteq K \subseteq G$. Show that H is also a normal subgroup of K.*

Solution. H is a normal subgroup of G. Therefore H is a subgroup of G. But K is a subgroup of G and $H \subseteq K$. Therefore H is also a subgroup of K. Now to show that H is normal in K. Let x be any element of K. Then $x \in K \Rightarrow x \in G$. Since H is normal in G, therefore we have $Hx=xH$. Thus H is a subgroup of K and we have $Hx=xH \ \forall \ x \in K$. Hence H is a normal subgroup of K.

Ex. 5. *If H is a subgroup of G and N is a normal subgroup of G, show that $H \cap N$ is a normal subgroup of H.*

(Meerut 1979, 81; Punjab 66; Sagar 68)

Solution. Since H and N are subgroups of G, therefore $H \cap N$ is also a subgroup of G. Also we have $H \cap N \subseteq H$. Therefore $H \cap N$ is a subgroup of H. Now to show that $H \cap N$ is a normal subgroup of H.

Let x be any element of H and a any element of $H \cap N$. Then $a \in H$ and $a \in N$. Since N is a normal subgroup of G, therefore we have $xax^{-1} \in N$. Also H is a subgroup of G. Therefore, we have $x \in H, a \in H \Rightarrow xax^{-1} \in H$.

Thus $xax^{-1} \in H \cap N$. Thus we have shown that

$$x \in H, a \in H \cap N \Rightarrow xax^{-1} \in H \cap N.$$

Consequently $H \cap N$ is a normal subgroup of H.

Ex. 6. *Show that a normal subgroup is commutative with every complex.* **(Meerut 1979)**

Solution. Suppose N is a normal subgroup and H is any complex of a group G. Then to prove that $HN=NH$.

Let $nh \in NH$ where $n \in N$, $h \in H$. We can write $nh=hh^{-1} nh$ $=h(h^{-1}nh)$. But N is a normal subgroup. Therefore $h^{-1}nh \in N$. Hence $nh \in HN$. Therefore $NH \subseteq HN$.

Again let $hn \in HN$ where $h \in H$, $n \in N$. We can write $hn=(hnh^{-1})h$ But $hnh^{-1} \in N$ because N is normal in G. Therefore $hn \in NH$. Thus $HN \subseteq NH$. Hence $HN=NH$.

Ex. 7. *If N is a normal subgroup of G and H is any subgroup of G, prove that NH is a subgroup of G.*

Solution. We know that a normal subgroup is commutative with every complex. [For proof see Ex. 6]

Since N is a normal subgroup of G and H is any subgroup of G, therefore we must have $NH=HN$. Now N and H are two subgroups of G such that $NH=HN$. Therefore NH is also a subgroup of G. [See Theorem 1 on page .44 of first chapter on groups].

Ex. 8. *If H is a subgroup of G and N is a normal subgroup of G, show that*

(*i*) *HN is a subgroup of G.*

(*ii*) *N is a normal subgroup of HN.* **(Punjab 1966)**

Solution. (i) Proceed as in Ex. 7.

(ii) Since $e \in H$, therefore obviously $N \subseteq HN$. Because if $n \in N$, then we can write $n = en$ which is an element of HN.

Now HN is a subgroup of G and N is also a subgroup of G. Also $N \subseteq HN$. Therefore N is a subgroup of HN. Now to show that N is a normal subgroup of HN. Let $h_1 n_1$ be any element of HN and n be any element of N. Then $h_1 \in H$, $n_1 \in N$ and we have $(h_1 n_1)\, n\, (h_1 n_1)^{-1} = h_1\, (n_1 n n_1^{-1})\, h_1^{-1} \in N$ since N is normal in G and $n_1 n n_1^{-1} \in N$, $h_1 \in G$. Therefore N is a normal subgroup of HN.

Ex. 9. *If N and M are normal subgroups of G, prove that NM is also a normal subgroup of G.* **(Nagarjuna 1978)**

Solution. We know that a normal subgroup is commutative with every complex. Therefore we have $NM = MN$. Now N and M are two subgroups of G such that $NM = MN$. Therefore NM is a subgroup of G.

Now to show that NM is a normal subgroup of G. Let x be any element of G and nm be any element of NM. Then $n \in N$, $m \in M$ and we have

$$x\,(nm)x^{-1} = (xn\,x^{-1})\,(xm\,x^{-1})$$

$$\in NM \qquad [\because\ N \text{ is normal} \Rightarrow xnx^{-1} \in N \text{ and } M \text{ is normal} \Rightarrow xmx^{-1} \in M].$$

Hence NM is a normal subgroup of G.

Ex. 10. *Suppose that N and M are two normal subgroups of G and that $N \cap M = \{e\}$. Show that every element of N commutes with every element of M.* **(Meerut 1971, 79, 81; Lucknow 70)**

Solution. Let n be any element of N and m any element of M. Then to prove that $nm = mn$. Consider the element $nmn^{-1}\,m^{-1}$.

Since N is normal, $mn^{-1}\,m^{-1} \in N$. Also $n \in N$. Therefore $nmn^{-1}\,m^{-1} \in N$.

Again M is normal $\Rightarrow nmn^{-1} \in M$. Also $m^{-1} \in M$. Therefore $nmn^{-1}\,m^{-1} \in M$. Thus

$$nmn^{-1}\,m^{-1} \in N \text{ and } nmn^{-1}\,m^{-1} \in M$$

$$\Rightarrow nmn^{-1}\,m^{-1} \in N \cap M$$

$$\Rightarrow nmn^{-1}\,m^{-1} = e \qquad [\because\ \{e\} = N \cap M]$$

$$\Rightarrow nm = mn.$$

Ex. 11. *Let G be a group, H a subgroup of G. Let, for*

$$g \in G,\ gHg^{-1}=\{ghg^{-1} : h\in H\}.$$

Prove that gHg^{-1} is a subgroup of G. **(Meerut 1979)**

Solution. Let gh_1g^{-1}, gh_2g^{-1} be any two elements of gHg^{-1}. Then $h_1, h_2 \in H$. Also we have

$$(gh_1 g^{-1})(gh_2g^{-1})^{-1}=gh_1g^{-1}(g^{-1})^{-1}h_2^{-1}g^{-1}=gh_1g^{-1}gh_2^{-1}g^{-1}$$

$$=gh_1h_2^{-1}g^{-1}\in gHg^{-1} \text{ since } h_1h_2^{-1}\in H.$$

$\therefore$ gHg^{-1} is a subgroup of G.

Ex. 12. *Suppose H is the only subgroup of finite order m in the group G. Prove that H is a normal subgroup of G.*

(Gujrat 1970)

Solution. H is a subgroup of G and $o(H)=m$. If x is any element of G then as in Ex. 11, xHx^{-1} is also a subgroup of G. We shall show that $o(xHx^{-1})=m$. Let $H=\{h_1, h_2, \ldots, h_m\}$.

Then $xHx^{-1}=\{xh_1x^{-1}, xh_2x^{-1}, \ldots, xh_mx^{-1}\}$.

The number of distinct elements in xHx^{-1} is m because

$$xh_ix^{-1}=xh_jx^{-1} \Rightarrow h_i=h_j.$$

$\therefore$ $o(xHx^{-1})=o(H)=m$.

But H is the only subgroup of G of order m. Therefore we must have $xHx^{-1}=H\ \forall\ x\in G$. Hence H is a normal subgroup of G.

Ex. 13. *If H is a subgroup of G, let*

$$N(H)=\{g\in G : gHg^{-1}=H\}.$$

Show that

(i) $N(H)$ is a subgroup of G.

(ii) H is a normal subgroup of $N(H)$.

(iii) If H is a normal subgroup of the subgroup K in G, then $K \subseteq N(H)$ i.e., $N(H)$ is the largest subgroup of G in which H is normal.

(iv) H is normal in G if and only if $N(H)=G$.

(Poona 1973)

Solution. Let $a, b\in N(H)$.

Then by def. of $N(H)$ we have $aHa^{-1}=H$ and $bHb^{-1}=H$.

Now $\quad bHb^{-1}=H$

$$\Rightarrow b^{-1}(bHb^{-1})b=b^{-1}Hb \Rightarrow H=b^{-1}Hb.$$

We have $(ab^{-1})H(ab^{-1})^{-1}=ab^{-1}H ba^{-1}=a(b^{-1}Hb)a^{-1}$

$$=aHa^{-1} \qquad [\because H=b^{-1}Hb]$$

$$=H.$$

$$\therefore \quad ab^{-1} \in N(H).$$

Thus $\quad a, b \in N(H) \Rightarrow ab^{-1} \in N(H)$.

Hence $N(H)$ is a subgroup of G.

(ii) Let h be any element of H. Since $hHh^{-1}=H$, therefore $h \in N(H)$. Thus $H \subseteq N(H)$. Therefore H is a subgroup of $N(H)$. Now to show that H is normal in $N(H)$. Let x be any element of $N(H)$. Then by definition of $N(H)$, we have $xHx^{-1}=H$. Therefore H is a normal subgroup of $N(H)$.

(iii) Let $k \in K$.

Since H is a normal subgroup of K, therefore we have $kHk^{-1}=H \Rightarrow k \in N(H)$. Thus $k \in K \Rightarrow k \in N(H)$. Therefore $K \subseteq N(H)$.

(iv) Let H be normal in G. Let $x \in G$. Then

$xHx^{-1}=H$ [$\because$ H is normal in G]

$\Rightarrow x \in N(H)$.

Thus $x \in G \Rightarrow x \in N(H)$. Therefore $G \subseteq N(H)$. But $N(H) \subseteq G$. Therefore $G=N(H)$.

Conversely let $N(H)=G$. Then $x \in G \Rightarrow x \in N(H)$ $\Rightarrow xHx^{-1}=H$. Thus we have $xHx^{-1}=H \ \forall \ x \in G$. Therefore H is a normal subgroup of G.

Ex. 14. *If a cyclic subgroup N of G is normal in G, then show that every subgroup of N is normal in G.*

Solution. Let a be a generator of N. Let H be any subgroup of N. We know that every subgroup of a cyclic group is cyclic. Therefore H is cyclic and a^m will be a generator of H where m is the least positive integer such that $a^m \in H$.

Now to show that H is normal in G. Suppose x is any arbitrary element of G and h is any arbitrary element of H. Then $h=(a^m)^k$ where k is some integer.

We have $xhx^{-1}=x\,(a^m)^k\,x^{-1}=x\,(a^k)^m x^{-1}$

$=(xa^k x^{-1})\,(xa^k x^{-1})\ldots$to m factors$=(xa^k x^{-1})^m$. ...(1)

Now a is a generator of N. Therefore $a^k \in N$.

Also N is normal in G. Therefore, we have

$x \in G,\ a^k \in N \Rightarrow xa^k x^{-1} \in N$.

$\therefore$ $xa^k x^{-1}=a^s$ for some integer s because a is a generator of N. Then from (1), we get $xhx^{-1}=(a^s)^m=(a^m)^s \in H$ because a^m is a generator of H.

195

Thus $x \in G, h \in H \Rightarrow xhx^{-1} \in H$. Hence H is a normal subgroup of G.

Ex. 15. *Give an example to show that if H is a normal subgroup of G and K is a normal subgroup of H then K may not be a normal subgroup of G.* **(Allahabad 1980)**

Solution. Consider the following subgroups of S_4, the symmetric group of permutations on four symbols a, b, c, d :

$G=\{I, (abcd), (adcb), (ab)(cd), (ac)(bd), (ad)(bc), (ac), (bd)\}$,

$H=\{I, (ab)(cd), (ac)(bd), (ad)(bc)\}$,

and $K=\{I, (ab)(cd)\}$.

It can be easily seen that H is a subgroup of G and K is a subgroup of H. Further any subgroup of index two in a group is a normal subgroup of the group. Here the index of H in G *i.e.*,

$$[G : H]=o(G)/o(H)=8/4=2.$$

Therefore H is normal in G. Also $[H : K]=o(H)/o(K)=4/2=2$. Therefore K is also normal in H. But here K is not a normal subgroup of G as can be easily seen. Take the element $(abcd) \in G$ and the element $(ab)(cd) \in K$. We have

$$(abcd)(ab)(cd)(abcd)^{-1}=(abcd)(ab)(cd)(dcba)=(ad)(bc) \notin K.$$

Therefore K is not a normal subgroup of G.

Exercises

1. Show that a subgroup H of a group G is normal if and only if the set G/H of all its left cosets is closed under (complex) multiplication. **(Punjab 1970; Dibrugarh 76; Meerut 78)**

[**Hint** : Proceed as in Theorem 3 on page 189].

2. Determine the coset decompositions of the subgroup

$$H=\{I, (1\ 2)\}$$

in S_3 the permutation group of degree 3 ; and further show that H is not normal in S_3. **(Rajasthan 1977)**

3. Show that a subgroup of index 2 is always an invariant (normal) subgroup. Hence show that alternating group A_n is an invariant subgroup of S_n the symmetric group of degree n. **(Lucknow 1970)**

4. If M and N are two normal subgroups of G and $M \cap N=(e)$, then prove that for any $n \in N$, $m \in M$

$$nm=mn.$$

5. Let G be a group and N a subgroup of G. Prove that the following statements are equivalent :

(i) $gng^{-1} \in N$ for all $g \in G$, $n \in N$.

(ii) $gNg^{-1}=N$ for all $g \in G$.

(iii) Every left coset of N in G is a right coset of N in G.

(iv) Product of two right cosets of N in G is again a right coset of N in G. **(Meerut 1974)**

6. Show that if H and N are subgroups of a group G, and N is normal in G, then $H \cap N$ is normal in H. Show by an example that $H \cap N$ need not be normal in G. **(I.A.S. 1971)**

7. Give an example of each of the following :

(i) A sub-group H of some group G, which is not normal in G.

(ii) A non-trivial sub-group H of a non-abelian group G, which is normal in G. **(Guru Nanak 1975)**

§ 2. Conjugate elements.

Definition. **(Agra 1969; Banaras 61)**

If a, b be two elements of a group G, then b is said to be conjugate to a if there exists an element x $\in$ *G such that*

$$b = x^{-1} ax.$$

If $b = x^{-1} ax$, then b is also called the **transform** of a by x.

If b is conjugate to a then symbolically we shall write $b \sim a$ and this relation in G will be called the *relation of conjugacy.* Thus $b \sim a$ iff $b = x^{-1} ax$ for some $x \in G$.

Theorem 1. *The relation of conjugacy is an equivalence relation on G.* **(Vikram 1976; Banaras 61)**

Proof. Reflexivity. If a is any element of G, then we have

$$a = e^{-1} ae \Rightarrow a \sim a.$$

Thus $a \sim a \; \forall \; a \in G$. Therefore the relation is reflexive.

Symmetry. We have $a \sim b \Rightarrow a = x^{-1} bx$ for some $x \in G$ $\Rightarrow xax^{-1} = x (x^{-1} bx) x^{-1} \Rightarrow xax^{-1} = b \Rightarrow b = (x^{-1})^{-1} ax^{-1}$ where $x^{-1} \in G$ $\Rightarrow b \sim a$.

Therefore the relation is symmetric.

Transitivity. Let $a \sim b$, $b \sim c$. Then $a = x^{-1} bx$, $b = y^{-1} cy$ for some $x, y \in G$. From this we get

$$a = x^{-1} (y^{-1} cy) \, x \qquad [\because \; b = y^{-1} cy]$$
$$= (yx)^{-1} \, c \, (yx) \text{ where } yx \in G.$$

$\therefore$ $a \sim c$ and thus the relation is transitive. Hence the relation of conjugacy in a group G is an equivalence relation. Therefore it will partition G into disjoint equivalence classes called **classes of conjugate elements**. These classes will be such that

(i) any two elements of the same class are conjugate.

(ii) no two elements of different classes are conjugate.

The collection of all elements conjugate to an element $a \in G$ will be symbolically denoted by $C(a)$ or by $\bar{a}$. Thus

$$C(a)=\{x\in G : x\sim a\}.$$

$C(a)$ will be called the *conjugate class of a in G*. We have $(y^{-1}ay)\sim a$ for all $y\in G$. Also if $b\sim a$ then b must be equal to $y^{-1'}ay$ for some $y\in G$. Therefore $C(a)=\{y^{-1}ay : y\in G\}$.

If G is a finite group, then the number of distinct elements in $C(a)$ will be denoted by c_a.

Normalizer of an element of a group.

Definition. *If $a\in G$, then $N(a)$, the normalizer of a in G is the set of all those elements of G which commute with a. Symbolically $N(a)=\{x\in G : ax=xa\}$.*

(I.A.S. 1975; Nagarjuna 78; Meerut 81, 84)

Theorem 2. *The normalizer $N(a)$ of $a\in G$ is a subgroup of G.*

(I.A.S. 1972; Nagpur 70; Delhi 70; Meerut 84; Punjab 70)

Proof. We have $N(a)=\{x\in G : ax=xa\}$.

Let $x_1, x_2\in N(a)$. Then $ax_1=x_1a$, $ax_2=x_2a$.

First we show that $x_2^{-1}\in N(a)$.

We have $ax_2=x_2a \Rightarrow x_2^{-1}(ax_2)x_2^{-1}=x_2^{-1}(x_2a)x_2^{-1}$

$$\Rightarrow x_2^{-1}a=ax_2^{-1} \Rightarrow x_2^{-1}\in N(a).$$

Now we shall show that $x_1x_2^{-1}\in N(a)$.

We have $a(x_1x_2^{-1})=(ax_1)x_2^{-1}=(x_1a)x_2^{-1}$

$$=x_1(ax_2^{-1})=x_1(x_2^{-1}a)=(x_1x_2^{-1})a.$$

$\therefore \quad x_1x_2^{-1}\in N(a)$.

Thus $x_1, x_2\in N(a) \Rightarrow x_1x_2^{-1}\in N(a)$.

$\therefore \quad N(a)$ is a subgroup of G.

Note 1. It should be noted that $N(a)$ is not necessarily a normal subgroup of G.

Note 2. Since $ex=xe\ \forall\ x\in G$, therefore $N(e)=G$.

Note 3. If G is an abelian group and $a\in G$, then

$xa=ax\ \forall\ x\in G$. Therefore $N(a)=G$.

Theorem 3. *Let a be any element of a group G. Then two elements x, y ∈ G give rise to the same conjugate of a if and only if they belong to the same right coset of the normalizer of a in G.*

Hence show that if G is a finite group, then $c_a=\dfrac{o(G)}{o[N(a)]}$, *i.e., the number of elements conjugate to a in G is the index of the normalizer of a in G.* **(Nagarjuna 1978; I.A.S. 72; Meerut 74)**

Proof. We have

$x, y\in G$ are in the same right coset of $N(a)$ in G

$\Leftrightarrow N(a)x=N(a)y$ [$\because$ $x \in N(a)x$, $y \in N(a)y$. Note that if H is a subgroup, then $x \in Hx$.]

$\Leftrightarrow xy^{-1} \in N(a)$ [$\because$ if H is a subgroup, then $Ha=Hb \Leftrightarrow ab^{-1} \in H$]

$\Leftrightarrow axy^{-1}=xy^{-1}a$ [by def. of $N(a)$]

$\Leftrightarrow x^{-1}(axy^{-1})\,y=x^{-1}(xy^{-1}a)\,y$

$\Leftrightarrow x^{-1}ax=y^{-1}ay$

$\Leftrightarrow$ x, y give rise to the same conjugate of a.

Hence the first result follows.

Now consider the right coset decomposition of G with respect to the subgroup $N(a)$. We have just proved that if $x, y \in G$ are in the same right coset of $N(a)$ in G, then they give the same conjugate of a. Further if x, y are in different right cosets of $N(a)$ in G, then they give rise to different conjugates of a. The reason is that if x, y give the same conjugate of a, then they must belong to the same right coset of $N(a)$ in G. Thus **there is a one-to-one correspondence between the right cosets of $N(a)$ in G and the conjugates of a.** So if G is a finite group, then

c_a = the number of distinct elements in $C(a)$

= the number of distinct right cosets of $N(a)$ in G

= the index of $N(a)$ in $G = \dfrac{o(G)}{o[N(a)]}$.

Corollary. *If G is a finite group, then*

$$o(G)=\Sigma\frac{o(G)}{o[N(a)]}$$

where this sum runs over one element a in each conjugate class.

(Punjab 1970 ; Meerut 84 P)

Proof. We know that the relation of conjugacy is an equivalence relation on G. Therefore it partitions G into disjoint conjugate classes. The union of all distinct conjugate classes will be equal to G and two distinct conjugate classes will have no common element. Since G is a finite group, therefore the number of distinct conjugate classes of G will be finite, say equal to k. Suppose $C(a)$ denotes the conjugate class of a in G and c_a denotes the number of elements in this class. If $C(a_1)$, $C(a_2)$,......, $C(a_k)$ are the k distinct conjugate classes of G, then

$$G=C(a_1) \cup C(a_2) \cup \ldots\ldots \cup C(a_k)$$

$\Rightarrow$ the number of elements in G = the number of elements in

$C(a_1)$+the number of elements in $C(a_2)$+……+the number of elements in $C(a_k)$.

[$\because$ two distinct conjugate classes have no common element]

$\Rightarrow o(G)=\Sigma\, c_a$, the summation being run over one element a in each conjugate class

$$\Rightarrow o\ (G)=\Sigma \frac{o\ (G)}{o\ [N\ (a)]} \text{ by previous theorem.}$$

Note. The equation in this corollary is often called the **class equation of** G.

Self-conjugate elements.

Definition. *An element* $a \in G$ *is said to be self-conjugate if* a *is the only member of the class* $C(a)$ *of elements conjugate to* a *i.e., if* $C(a)=\{a\}$.

Thus, a, is self-conjugate if and only if $a=x^{-1}ax\ \forall\ x \in G$

or $xa=ax\ \forall\ x \in G.$

Thus a self-conjugate element is one which commutes with each element of the group. If a is a self-conjugate element, then we have $a=x^{-1}ax\ \forall\ x \in G$.

Thus the transform of a by every element of G remains equal to a. Therefore sometimes a self-conjugate element is also called an **invariant element.**

The centre of a group.

Definition. *The set* Z *of all self-conjugate elements of a group* G *is called the centre of* G. *Symbolically*

$Z=\{z \in G : zx=xz\ \forall\ x \in G\}$. **(Punjab 1968; B.H.U. 71)**

Theorem 4. *The centre* Z *of a group* G *is a normal subgroup of* G. **(Banaras 1971; Meerut 81)**

Proof. We have $Z=\{z \in G : zx=xz\ \forall\ x \in G\}$.

First we shall prove that Z is a subgroup of G.

Let $z_1, z_2 \in Z$. Then $z_1x=xz_1$ and $z_2x=xz_2$ for all $x \in G$.

We have $z_2x=xz_2\ \forall\ x \in G$

$$\Rightarrow z_2^{-1}(z_2x)\,z_2^{-1}=z_2^{-1}(xz_2)\,z_2^{-1}$$

$$\Rightarrow xz_2^{-1}=z_2^{-1}x\ \forall\ x \in G$$

$$\Rightarrow z_2^{-1} \in Z.$$

Now $(z_1z_2^{-1})\,x=z_1\,(z_2^{-1}x)=z_1\,(xz_2^{-1})=(z_1x)\,z_2^{-1}=(xz_1)\,z_2^{-1}$

$$=x\,(z_1z_2^{-1}).$$

$\therefore\ z_1z_2^{-1} \in Z.$

Thus $z_1, z_2 \in Z \Rightarrow z_1z_2^{-1} \in Z.$

$\therefore$ **Z is a subgroup of G.**

Now we shall show that Z is a normal subgroup of G. Let $x \in G$ and $z \in Z$. Then

$$xzx^{-1}=(xz)\,x^{-1}=(zx)\,x^{-1}=z \in Z.$$

Thus $\quad x \in G, z \in Z \Rightarrow xzx^{-1} \in Z.$

$\therefore$ Z is a normal subgroup of G.

Theorem 5. *$a \in Z$ if and only if $N(a)=G$. If G is finite, $a \in Z$ if and only if $o\,[N(a)]=o\,(G)$.* **(Nagarjuna 1978)**

Proof. Let $a \in Z$. Then by def. of Z, we have

$$ax=xa \;\forall\; x \in G.$$

Also $\quad N(a)=\{x \in G : ax=xa\}$.

Now $\quad a \in Z \Leftrightarrow ax=xa \;\forall\; x \in G$ [by def. of Z]

$\Leftrightarrow x \in N(a) \;\forall\; x \in G$ [by def. of $N(a)$]

$\Leftrightarrow N(a)=G$. [$\because$ $N(a) \subseteq G$ and each element of G is in $N(a)$]

If the group G is finite, then $N(a)=G \Leftrightarrow o\,(G)=o\,[N(a)]$.

Therefore if the group G is finite, then $a \in Z$ if and only if

$$o\,[N(a)]=o\,(G).$$

Theorem 6. *Let G be a finite group and Z be the centre of G. Then the class equation of G can be written as*

$$o\,(G)=o\,(Z)+\sum_{a \notin Z} \frac{o\,(G)}{o\,[N(a)]},$$

where the summation runs over one element a in each conjugate class containing more than one element.

Proof. The class equation of G is

$o\,(G)=\Sigma \dfrac{o\,(G)}{o\,[N(a)]}$, the summation being extended over one element a in each conjugate class.

Now $a \in Z \Leftrightarrow o\,[N(a)]=o\,(G) \Leftrightarrow o\,(G)/o\,[N(a)]=1 \Leftrightarrow$ the conjugate class of a in G contains only one element. Thus the number of conjugate classes each having only one element is equal to $o\,(Z)$. If a is an element of any one of these conjugate classes, we have $o\,(G)/o\,[N(a)]=1$. Hence the class equation of G takes the desired form

$$o\,(G)=o\,(Z)+\sum_{a \notin Z} \frac{o\,(G)}{o\,[N(a)]}.$$

Theorem 7. *If $o\,(G)=p^n$ where p is a prime number, then the centre $Z \neq \{e\}$.* **(I.A.S. 1972 ; Punjab 78 ; Meerut 74)**

Proof. By the class equation of G, we have

$$o(G)=o(Z)+\sum_{a\notin Z}\frac{o(G)}{o[N(a)]}, \qquad \text{...(1)}$$

where the summation runs over one element a in each conjugate class containing more than one element.

Now $\forall\ a\in G$, $N(a)$ is a subgroup of G. Therefore by Lagrange's theorem, $o[N(a)]$ is a divisor of $o(G)$. Also $a\notin Z \Rightarrow N(a)\neq G \Rightarrow o[N(a)]<o(G)$. Therefore if $a\notin Z$, then $o[N(a)]$ must be of the form p^{n_a} where n_a is some integer such that $1\leqslant n_a<n$. Suppose there are exactly z elements in Z *i.e.*, let $o(Z)=z$. Then the class equation (1) gives

$$p^n=z+\Sigma\frac{p^n}{p^{n_a}}, \text{ where each } n_a \text{ is some integer such that } 1\leqslant n_a<n.$$

$$z=p^n-\Sigma\frac{p^n}{p^{n_a}}, \qquad \text{...(2)}$$

where n_a's are some positive integers each being less than n.

Now $p\mid p^n$. Also p divides each term in the Σ of the right hand side of (2) because each $n_a<n$. Thus we see that p is a divisor of the right hand side of (2). Therefore p is a divisor of z. Now $e\in Z$. Therefore $z\neq 0$. Therefore z is a positive integer divisible by the prime p. Therefore $z>1$. Hence Z must contain an element besides e. Therefore $Z\neq\{e\}$.

Corollary. *If $o(G)=p^2$ where p is a prime number, then G is abelian.* **(Kumayun 1977 ; Kanpur 80 ; Meerut 81)**

Proof. We shall show that the centre Z of G is equal to G itself. Then obviously G will be an abelian group.

Since p is a prime number, therefore by the previous theorem $Z\neq\{e\}$. Therefore $o(Z)>1$. But Z is a subgroup of G, therefore $o(Z)$ must be a divisor of $o(G)$ *i.e.*, $o(Z)$ must be a divisor of p^2. Since p is prime, therefore either $o(Z)=p$ or p^2.

If $o(Z)=p^2$, then $Z=G$ and our proof is complete.

Now suppose that $o(Z)=p$. Then $o(Z)<o(G)$ because $p<p^2$. Therefore there must be an element which is in G but which is not in Z. Let $a\in G$ and $a\notin Z$.

Now $N(a)$ is a subgroup of G and $a\in N(a)$. Also $x\in Z \Rightarrow xa=ax$ and this implies $x\in N(a)$. Thus $Z\subseteq N(a)$. Since $a\notin Z$, therefore the number of elements in $N(a)$ is $\not> p$ *i.e.*, $o[N(a)]>p$. But order of $N(a)$ must be a divisor of p^2. Therefore

$o\,[N\,(a)]$ must be equal to p^2. Then $N\,(a)=G$. Therefore $a\in Z$ and thus we get a contradiction.

Therefore it is not possible that $o\,(Z)=p$. Hence the only possibility is that

$$o\,(Z)=p^2 \Rightarrow Z=G \Rightarrow G \text{ is abelian.}$$

Ex. *Is a group of order* 121 *abelian ?* **(Meerut 1977)**

Ans. Yes.

§ 3. Conjugate subgroups :

Definition. *If A, B be two subgroups of a group G, then B is said to be conjugate to A if there exists an element $x \in G$ such that*

$$B=x^{-1}\,Ax.$$

If $B=x^{-1}\,Ax$, then B is also called the *transform* of A by x.

If B is conjugate to A, then symbolically we shall write $B \sim A$.

Theorem 1. *The relation of being conjugate is an equivalence relation on the set of subgroups of a group G.*

Proof. Proceed as in theorem 1 of § 2, page 197.

The relation of conjugacy in the family of subgroups of a group G will partition the family into disjoint equivalence classes. The collection of all subgroups conjugate to a subgroup A of G will be symbolically denoted by $C\,(A)$. Obviously

$$C\,(A)=\{x^{-1}\,Ax : x\in G\}.$$

Normalizer of a subgroup of a group.

Definition. *If A is a subgroup of a group G, then $N\,(A)$, the normalizer of A in G is the set of all those elements of G which commute with A. Symbolically $N\,(A)=\{x\in G : xA=Ax\}$.*

Theorem 2. *The normalizer $N\,(A)$ of a subgroup A of a group G is a subgroup of G.* **(Punjab 1969)**

Proof. Proceed as in Theorem 2 of § 2.

Theorem 3. *Suppose A is a subgroup of a group G. Then there is a one-to-one correspondence between the right cosets of $N\,(A)$ in G and the conjugates of A.* **(Punjab 1969)**

Proof. Proceed as in Theorem 3 of § 2.

Self-conjugate subgroups.

Definition. *A subgroup A of a group G is said to be self-conjugate if A is the only member of the class $C(A)$ of subgroups conjugate to A.*

Thus, A, is self conjugate iff

$$A=x^{-1}\,Ax \;\forall\; x\in G$$

or $xA=Ax \; \forall \; x \in G$

or A is a normal subgroup of G.

If A is a self-conjugate subgroup of a group G, then we have $A=x^{-1} Ax \; \forall \; x \in G$. Thus the transform of A by every element of G remains equal to A. Therefore sometimes a self-conjugate subgroup is also called an **invariant subgroup**. It is quite obvious that a subgroup of a group G is invariant if and only if it is normal. Therefore sometimes a normal subgroup is also called an invariant subgroup.

§ 4. **Quotient Groups.** We are now going to introduce the very important concept of *Quotient Groups*.

Let H be any normal subgroup of a group G. If $a \in G$, then Ha is a right coset of H in G. Also H being normal in G, the left coset aH will be equal to the right coset Ha. Thus there is no distinction between right and left cosets. So we can say that Ha is a coset of H in G. Let G/H be the collection of all cosets of H in G *i.e.*, let $G/H=\{Ha : a \in G\}$.

If $a, b \in G$, then we have

$$(Ha)(Hb)=H(aH)b$$
$$=H(Ha)b \qquad [\because Ha=aH \text{ as } H \text{ is normal}]$$
$$=HHab=Hab. \qquad [\because HH=H]$$

Now $ab \in G$, therefore the product of two cosets of H in G is again a coset of H in G. We shall presently see that the set G/H is a group with respect to multiplication of cosets.

Theorem. *The set of all cosets of a normal subgroup is a group with respect to multiplication of complexes as the composition.*

(Meerut 1979; Kanpur 80; Nagarjuna 80; Rajasthan 78)

Proof. Let H be a normal subgroup of a group G. Since H is normal in G, therefore each right coset will be equal to the corresponding left coset. Thus there is no distinction between right and left cosets and we shall call them simply as cosets. Let G/H be the collection of all cosets of H in G *i.e.*, let $G/H=\{Ha : a \in G\}$.

Closure Property. Let $a, b \in G$. Then

$$(Ha)(Hb)=H(aH)b=H(Ha)b=HHab=Hab.$$

Since $ab \in G$, therefore Hab is also a coset of H in G. So $Hab \in G/H$. Thus G/H is closed with respect to coset multiplication.

Associativity. Let $a, b, c \in G$. Then $Ha, Hb, Hc \in G/H$. We have $Ha\,[(Hb)(Hc)]=Ha\,(Hbc)=Ha\,(bc)$

$=H\,(ab)\,c$ $\quad [\because \ a\,(bc)=(ab)c]$

$=(Hab)\,Hc=[(Ha)\,(Hb)]\,Hc.$

Thus the product in G/H satisfies the associative law.

Existence of Identity. We have $H=He\in G/H$. Also if Ha is any element of G/H, then

$$H\,(Ha)=(He)\,(Ha)=Hea=Ha \text{ and similarly}$$

$$(Ha)\,H=(Ha)\,(He)=Hae=Ha.$$

Therefore the coset H is the identity element.

Existence of Inverse. Let $Ha\in G/H$. Then $Ha^{-1}\in G/H$.

We have $\quad (Ha)\,(Ha^{-1})=Haa^{-1}=He=H$

and $\quad (Ha^{-1})\,(Ha)=Ha^{-1}a=He=H.$

$\therefore$ The coset Ha^{-1} is the inverse of Ha *i.e.*, $(Ha)^{-1}=Ha^{-1}$. Thus each element of G/H possesses inverse.

Hence G/H is a group with respect to product of cosets.

Definition. Quotient Group.

(I.A.S. 1970; Meerut 79; Garhwal 76; Mysore 70)

If G is a group and H is a normal subgroup of G, then the set G/H of all cosets of H in G is a group with respect to multiplication of cosets. It is called the **quotient group or factor group** *of G by H.*

The identity element of the quotient group G/H is H.

Solved Examples

Ex. 1. *Let* **I** *be the additive group of integers. Let H be a subgroup of* **I** *such that* $H=\{mx : x\in \mathbf{I}\}$ *where m is a fixed positive integer. Write the elements of the quotient group* **I**/*H. Also prepare a composition table for* **I**/*H when* $m=5$. **(Meerut 1976)**

Solution. Since **I** is an abelian group, therefore H is normal in **I**. The elements of **I**/H are the cosets of H in **I** namely

$H+0=H=\{\ldots -2m,\ -m,\ 0,\ m,\ 2m,\ \ldots\}$

$H+1=\{\ldots,\ -2m+1,\ -m+1,\ 1,\ m+1,\ 2m+1,\ldots\}$

$H+2=\{\ldots,\ -2m+2,\ -m+2,\ 2,\ m+2,\ 2m+2,\ldots\}$

... ...

... ...

$H+(m-2)=\{\ldots,\ -m-2,\ -2\ \ m-2,\ 2m-2,\ 3m-2,\ \ldots\}$

$H+(m-1)=\{\ldots,\ -m-1,\ -1,\ m-1,\ 2m-1,\ 3m-1,\ \ldots\}.$

These are the only distinct cosets of H in **I**. Because if s is any integer, then by division algorithm there exist integers q and r such that $\quad s=mq+r$ where $0\leqslant r\leqslant m-1$.

We have $\quad H+s=H+mq+r$

$=H+r \quad [\because \ mq\in H$ and this gives $H+mq=H]$.

Thus $H+s$ is one of the above m cosets of H in I. Thus there are m distinct elements in the set $\mathbf{I}/H$.

When $m=5$, the distinct elements in $\mathbf{I}/H$ are

$$H,\ H+1,\ H+2,\ H+3,\ H+4.$$

If $a, b \in \mathbf{I}$, then $(H+a)+(H+b)=H+(a+b)$.

Also $H+a=H+b \Leftrightarrow a-b \in H$. Thus $H+2=H+7$, $H+3=H+8$ and so on.

Hence the composition table for $\mathbf{I}/H$ is as given below :

	H	$H+1$	$H+2$	$H+3$	$H+4$
H	H	$H+1$	$H+2$	$H+3$	$H+4$
$H+1$	$H+1$	$H+2$	$H+3$	$H+4$	H
$H+2$	$H+2$	$H+3$	$H+4$	H	$H+1$
$H+3$	$H+3$	$H+4$	H	$H+1$	$H+2$
$H+4$	$H+4$	H	$H+1$	$H+2$	$H+3$

Ex. 2. *Let P_3 be the symmetric group on three symbols a, b c and A_3 be the alternating group on three symbols a, b, c. Form the composition table for the quotient group P_3/A_3.* **(Rajasthan 1977)**

Solution. Let $P_3=\{f_1, f_2, f_3, f_4, f_5, f_6\}$ where f_1=identity permutation, $f_2=(ab)$, $f_3=(bc)$, $f_4=(ca)$, $f_5=(abc)$, $f_6=(acb)$.

We have A_3=set of even permutations belonging to P_3. Thus $A_3=\{f_1, f_5, f_6\}$.

A_3 is a normal subgroup of P_3 as we have already proved.

The elements of $P_3/\;_3$ are the cosets of A_3 in P_3. By Lagrange's theorem A_3 will have only two distinct cosets in P_3. One is A_3 itself and the other is $A_3 f_2$. Thus A_3, $A_3 f_2$ are the only two elements of P_3/A_3. It should be noted that

$$A_3 f_5=A_3 f_6=A_3,\ A_3 f_2=A_3 f_3=A_3 f_4.$$

The composition table for P_3/A_3 is as given below :

	A_3	$A_3 f_2$
A_3	A_3	$A_3 f_2$
$A_3 f_2$	$A_3 f_2$	A_3

[Note that $(A_3 f_2)\ (A_3 f_2)]$
$=A_3 f_2 f_2=A_3 f_1=A_3$

Ex. 3. *If G is a finite group and H is a normal subgroup of G, then $o\,(G/H)=\dfrac{o\,(G)}{o\,(H)}$.* (Meerut 1977, 80, 82; Kumayun 77)

Solution. We have

$$o\,(G/H)=\text{number of distinct right cosets of } H \text{ in } G$$
$$=\frac{\text{Number of elements in } G}{\text{Number of elements in } H} \quad \text{[by Lagrange's theorem]}$$
$$=\frac{o\,(G)}{o\,(H)}.$$

Ex. 4. *Show that every quotient group of an abelian group is abelian and the converse is not true.* (Rajasthan 1974)·

Solution. Let G be an abelian group and H be a subgroup of G. Then H is a normal subgroup of G. If $a, b \in G$, then Ha, Hb are any two elements of G/H. We have

$$(Ha)\,(Hb)=Hab=Hba \qquad [\quad G \text{ is abelian} \Rightarrow ab=ba]$$
$$=(Hb)\,(Ha).$$

$\therefore$ G/H is abelian.

The converse is not true. For example if P_3 be the symmetric group of degree 3 and A_3 be the alternating group of degree 3, then P_3/A_3 is an abelian group while P_3 is not an abelian group. The group P_3/A_3 is of order 2, and every group of order 2 is abelian.

Ex. 5. *If N is normal in G and $a \in G$ is of order n, prove that the order, m, of Na in G/N is a divisor of n.*

Solution. The identity of the quotient group G/N is N. If e is the identity of G, then $\quad o(a)=n \Rightarrow a^n=e.$

$\therefore$ $Na^n=Ne=N.$

But $\quad Na^n=N(aaa\ldots\text{upto } n \text{ times})$
$\quad\quad =(Na)\,(Na)\,(Na)\ldots\text{upto } n \text{ times}=(Na)^n.$

Thus $\quad (Na)^n=N$ (identity of G/N).

Since order of Na in G/N is m and $(Na)^n=N$, identity of G/N, therefore m must be a divisor of n. [Refer Theorem 4, page 115].

Ex. 6. *Show that every quotient group of a cyclic group is cyclic and the converse is not true.* (Jodhpur 1970; Meerut 78)

Solution. Let G be a cyclic group and a be a generator of G. Let H be a subgroup of G. Since every cyclic group is abelian, therefore H is a normal subgroup of G. Let a^n be any element of

G where n is some integer. Then Ha^n is any element of G/H. As can be easily seen $Ha^n=(Ha)^n$ for every integer n. Therefore G/H is a cyclic group and Ha is a generator of it.

The converse is not true. For example P_3/A_3 is cyclic while P_3 is not cyclic.

Ex. 7. *Let Z be the centre of a group G. If $a \in Z$, then prove that the cyclic subgroup $\{a\}$ of G generated by a is a normal subgroup of G.*

Solution. We have

$Z=\{z \in G : zx=xz \ \forall \ x \in G\}$. Let $a \in Z$ and let $H=\{a\}$ be the cyclic subgroup of G generated by a. Let h be any element of H. Then $h=a^n$ for some integer n.

Let x be any element of G. We have

$$xhx^{-1}=xa^nx^{-1}$$

$$=(xax^{-1})^n \qquad \text{[See Ex. 3 page 117 of first chapter on groups]}$$

$$=(axx^{-1})^n \qquad [\because \ a \in Z \Rightarrow ax=xa]$$

$$=(ae)^n=a^n \in H.$$

Thus $xhx^{-1} \in H \ \forall \ h \in H$ and $\forall \ x \in G$.

$\therefore$ H is a normal subgroup of G.

Ex. 8. *Let a be any element of G. Show that the cyclic subgroup of G generated by a is a normal subgroup of the normalizer of a.* **(Punjab 1970)**

Solution. We have the normalizer of a

$$=N(a)=\{x \in G : xa=ax\}.$$

Let H be the cyclic subgroup of G generated by a. Let h be any element of H. Then $h=a^n$ where n is some integer. We have

$$a^na=a^{n+1}=aa^n.$$

$\therefore$ $a^n=h \in N(a)$.

Now $N(a)$ and H are subgroups of G. Also $h \in H \Rightarrow h \in N(a)$. Therefore $H \subseteq N(a)$. Hence H is a subgroup of $N(a)$.

Now to prove that H is a normal subgroup of $N(a)$. Let x be any element of $N(a)$ and $h=a^n$ be any element of H. We have

$$xhx^{-1}=xa^nx^{-1}=(xax^{-1})^n$$

$$=(axx^{-1})^n \qquad [\because \ x \in N(a) \Rightarrow ax=xa].$$

$$=(ae)^n=a^n \in H.$$

$\therefore$ H is a normal subgroup of $N(a)$.

Ex. 9. *Show that two elements are conjugate if and only if they can be put in the form xy and yx respectively where x and y are suitable elements of G.* **(Punjab 1970)**

Solution. Let a, b be two conjugate elements of a group G.

Then $a=c^{-1}bc$ for some $c \in G$.

Let $c^{-1}b=x$ and $c=y$. Then $a=xy$. Also

$$yx=c\,(c^{-1}b)=(cc^{-1})\,b=eb=b.$$

Conversely suppose that $a=xy$ and $b=yx$. We have

$$b=yx \Rightarrow y^{-1}b=y^{-1}yx \Rightarrow y^{-1}b=x.$$

Now $a=xy$

$\Rightarrow a=y^{-1}\,by \Rightarrow a$ and b are conjugate elements.

Ex. 10 *Give an example to show that in a group G the normaliser of an element is not necessarily a normal subgroup of G.*

Solution. Consider the group S_3, the symmetric group of permutations on three symbols a, b, c. We have $S_3=\{I, (ab), (bc), (ca), (abc), (acb)\}$. Let $N\,(ab)$ denote the normaliser of the element $(ab) \in S_3$. We shall show that $N\,(ab)$ is not a normal subgroup of S_3. Let us calculate the elements of $N(ab)$. Obviously $(ab) \in N\,(ab)$. Also $I \in N\,(ab)$ because $I\,(ab)=(ab)\,I$.

Now $(bc)\,(ab)=(abc)$, and $(ab)\,(bc)=(acb)$. Thus (bc) does not commute with (ab). Therefore $(bc) \notin N\,(ab)$. Again

$$(ca)\,(ab)=(acb) \text{ and } (ab)\,(ca)=(abc).$$

Thus $(ca)\,(ab) \neq (ab)\,(ca)$ and therefore $(ca) \notin N(ab)$. Similarly we can verify that $(abc) \notin N\,(ab)$ and $(acb) \notin N\,(ab)$. Hence

$$N\,(ab)=\{I, (ab)\}.$$

Now we shall show that $N\,(ab)$ is not a normal subgroup of S_3. Take the element $(bc) \in S_3$ and the element $(ab) \in N\,(ab)$. We have $(bc)\,(ab)\,(bc)^{-1}=(bc)\,(ab)\,(cb)=(abc)\,(cb)=(ac) \notin N\,(ab)$. Therefore $N\,(ab)$ is not a normal subgroup of S_3.

Ex. 11. *Let N_1 and N_2 be two normal subgroups of a group G. Prove that $G/N_1=G/N_2$ if and only if $N_1=N_2$.*

Solution. If $N_1=N_2$, then obviously $G/N_1=G/N_2$.

Conversely suppose that $G/N_1=G/N_2$. Then we are to prove that $N_1=N_2$. We have $N_1 \in G/N_1$. But $G/N_1=G/N_2$. Therefore $N_1 \in G/N_2$ *i.e.*, N_1 is equal to some coset of N_2 in G. But two cosets of N_2 in G are either disjoint or identical. Since $e \in N_1$ and $e \in N_2$, therefore N_1 and N_2 are not disjoint. So we must have

$$N_1=N_2.$$

Ex. 12. *Let Z denote the centre of a group G. If G/Z is cyclic prove that G is abelian.* **(Meerut 1978, 81)**

Solution. It is given that G/Z is cyclic. Let Zg be a generator of the cyclic group G/Z where g is some element of G.

Let $a, b \in G$. Then to prove that $ab=ba$. Since $a \in G$, therefore $Za \in G/Z$. But G/Z is cyclic having Zg as a generator. Therefore there exists some integer m such that $Za=(Zg)^m=Zg^m$, because Z is a normal subgroup of G. Now $a \in Za$. Therefore

$$Za=Zg^m \Rightarrow a \in Zg^m \Rightarrow a=z_1 g^m \text{ for some } z_1 \in Z.$$

Similarly $b=z_2 g^n$ where $z_2 \in Z$ and n is some integer.

$$
\begin{aligned}
\text{Now} \quad ab &= (z_1 g^m)(z_2 g^n) = z_1 g^m z_2 g^n \\
&= z_1 z_2 g^m g^n \qquad [\because z_2 \in Z \Rightarrow z_2 g^m = g^m z_2] \\
&= z_1 z_2 g^{m+n}.
\end{aligned}
$$

$$
\begin{aligned}
\text{Again} \quad ba &= z_2 g^n z_1 g^m = z_2 z_1 g^n g^m = z_2 z_1 g^{n+m} \\
&= z_1 z_2 g^{m+n} \qquad [\because z_1 \in Z \Rightarrow z_1 z_2 = z_2 z_1]
\end{aligned}
$$

$$\therefore \quad ab=ba.$$

Since $ab=ba \ \forall \ a, b \in G$, therefore G is abelian.

Ex. 13. *If p is a prime number and G is a non-abelian group of order p^3, show that the centre of G has exactly p elements.* **(Madras 1983)**

Solution Let Z denote the centre of G. Since $o(G)=p^3$ where p is a prime number, therefore $Z \neq \{e\}$ *i.e.*, $o(Z) > 1$. But Z is a subgroup of G, therefore $o(Z)$ must of a divisor of $o(G)$ *i.e* $o(Z)$ must be a divisor of p^3. Since p is prime, therefore either $o(Z)=p$ or p^2 or p^3.

If $o(Z)=p^3=o(G)$, then $Z=G$ and so G is abelian which contradicts the hypothesis that G is non-abelian. So $o(Z)$ cannot be p^3.

If $o(Z)=p^2$, then $o(G/Z)=o(G)/o(Z)=p^3/p^2=p$ *i.e.*, G/Z is a group of prime order p and so is cyclic. But if G/Z is cyclic, then G is abelian which again contradicts the hypothesis. So $o(Z)$ cannot be p^2.

Hence the only possibility is that $o(Z)=p$ *i.e.*, the centre of G has exactly p elements.

§ 5. Homomorphisms of Groups.

Definition. Homomorphism into. *A mapping f from a group G into a group G' is said to be a homomorphism of G into G' if*

$$f(ab)=f(a)\,f(b)\ \forall\ a, b\in G.$$

Homomorphism onto *A mapping f from a group G onto a group G' is said to be a homomorphism of G onto G' if*

$$f(ab)=f(a)\,f(b)\ \forall\ a, b\in G.$$

Also then G' is said to be a homomorphic image of G.

(I.A.S 1970; Kanpur 80; Gujrat 71; Allahabad 79; Banaras 62; Sambalpur 77; Rajasthan 77; Madras 83; Meerut 70)

It should be noted that isomorphism is a special type of homomorphism. If f is a homomorphism of G into G' and f is one-one, then f is an isomorphism of G into G'. Similarly if f is a homomorphism of G onto G' and f is one-one, then f is an isomorphism of G onto G'.

Endomorphism. *A homomorphism of a group into itself is called an endomorphism.* **(Andhra 1975; Rajasthan 77)**

Example 1 *Show that the mapping f of the symmetric group P_n onto the multiplicative group $G'=\{1, -1\}$ defined by*

$$f(\alpha)=1 \quad or \quad -1$$

according as α is an even or odd permutation in P_n is a homomorphism of P_n onto G'.

Solution. We know that the product of two permutations both even or both odd is even while the product of one even or one odd permutation is odd. We shall show that

$$f(\alpha\beta)=f(\alpha)\,f(\beta)\ \forall\ \alpha, \beta\in P_n.$$

(i) If α, β are both even, then
$f(\alpha\beta)=1=1.1=f(\alpha)\,f(\beta)$.

(ii) If α, β are both odd, then
$f(\alpha\beta)=1=(-1)(-1)=f(\alpha)\,f(\beta)$.

(iii) If α is odd and β is even, then
$f(\alpha\beta)=-1=(-1)(1)=f(\alpha)\,f(\beta)$.

(iv) If α is even and β is odd, then
$f(\alpha\beta)=-1=(1)(-1)=f(\alpha)\,f(\beta)$.

Thus $f(\alpha\beta)=f(\alpha)\,f(\beta)\ \forall\ \alpha, \beta\in P_n$.

Also obviously f is onto G'. Therefore f is a homomorphism of P_n onto G'.

Example 2. *Let G be the group of all ordered pairs* (a, b) *of real numbers with the binary operation denoted additively and defined by*

$$(a, b)+(c, d)=(a+c, b+d).$$

Further let G' be the additive group of all real numbers. Then the mapping $f: G \to G'$ *defined by* $f(a, b)=a \ \forall \ (a, b) \in G$ *is a homomorphism of G onto G'.*

Solution. It can be easily proved that G is a group with respect to the given binary operation. The ordered pair $(0, 0)$ is the identity element and the ordered pair $(-a, -b)$ is the inverse of (a, b).

Let (a, b) and (c, d) be any two elements of G.

Then by definition of f, we have

$$f(a, b)=a, f(c, d)=c.$$

Now $f[(a, b)+(c, d)]=f(a+c, b+d)=a+c=f(a, b)+f(c, d)$.

Also obviously f is onto G'. Therefore f is a homomorphism of G onto G'.

Example 3. *Let G be a group and let e be the identity element of G. Then the mapping* $f: G \to G$ *defined by* $f(a)=e \ \forall \ a \in G$ *is an endomorphism of G.* **(Sambalpur 1977)**

Solution. Let a, b be any two elements of G. Then

$$f(a)=e, f(b)=e.$$

Now $f(ab)=e=ee=f(a)f(b)$.

Thus f is a homomorphism of G into G. Therefore f is an endomorphism of G.

Theorem. *If f is a homomorphism of a group G into a group G', then*

(*i*) $f(e)=e'$, *where e is the identity of G and e' is the identity of G'.* **(Gujrât 1970; Rajasthan 76; Banaras 70; Meerut 79)**

(*ii*) $f(a^{-1})=[f(a)]^{-1} \ \forall \ a \in G$.
(Gujrat 1970; Kanpur 80; Banaras 70; Meerut 79)

(*iii*) *If the order of* $a \in G$ *is finite, then the order of* $f(a)$ *is a divisor of the order of a.*

Proof. (i) Let $a \in G$. Then $f(a) \in G'$. We have

$$\begin{aligned} f(a)\, e' &= f(a) && [\because\ e' \text{ is the identity of } G'] \\ &= f(ae) && [\because\ e \text{ is the identity of } G] \\ &= f(a)\, f(e) && [\because\ f \text{ is a homomorphism}] \end{aligned}$$

Now G' is a group. Therefore

$$f(a)\, e' = f(a)\, f(e)$$
$$\Rightarrow e' = f(e) \qquad [\text{by left cancellation law in } G']$$

(ii) Let a be any element of G. Then $a^{-1} \in G$.

We have $e' = f(e) = f(aa^{-1}) = f(a)\, f(a^{-1})$.

Therefore $f(a^{-1})$ is the inverse of $f(a)$ in the group G'. Thus

$$f(a^{-1}) = [f(a)]^{-1}.$$

(iii) Let $a \in G$ and $o(a) = m$. We have $o(a) = m \Rightarrow a^m = e$.

$\therefore\ f(a^m) = f(e)$

$\Rightarrow f(aaa\ldots m \text{ times}) = e'$

$\Rightarrow f(a)\, f(a) \ldots m \text{ times} = e' \Rightarrow [f(a)]^m = e'$.

$\therefore$ if n is the order of $f(a)$ in G', then n must be a divisor of m. [Refer theorem 4, page 115].

§ 6. Kernel of a Homomorphism.

Definition. *If f is a homomorphism of a group G into a group G', then the set K of all those elements of G which are mapped by f onto the identity e' of G' is called the kernel of the homomorphism f.*

(Kanpur 1969; Rajasthan 78; Meerut 74)

Thus if f is a homomorphism of G into G', then K is the kernel of f if $K = \{x \in G : f(x) = e'$ where e' is the identity of $G'\}$.

Theorem 1. *If f is a homomorphism of a group G into a group G' with kernel K, then K is a normal subgroup of G.*

(G. N. D. U. Amritsar 1982; Kumayun 77; Bombay 70; Madras 83; Marathwada 72; Meerut 79, 81, 83, 84; Rajasthan 78)

Proof. Let f be a homomorphism of a group G into a group G'. Let e, e' be the identities of G and G' respectively. Let K be the kernel of f. Then $K = \{x \in G : f(x) = e'\}$.

Since $f(e) = e'$, therefore at least $e \in K$. Thus K is not empty.

Let $a, b \in K$. Then $f(a) = e'$, $f(b) = e'$. We have

$$f(ab^{-1}) = f(a)\, f(b^{-1}) = f(a)\, [f(b)]^{-1} = e' e'^{-1} = e' e' = e'.$$
$$\therefore\ ab^{-1} \in K.$$

Thus $\qquad a, b \in K \Rightarrow ab^{-1} \in K.$

Therefore K is a subgroup of G. Now to prove that K is normal in G. Let g be any element of G and k be any element of K. Then $f(k) = e'$. We have

$$f(gkg^{-1})=f(g)f(k)f(g^{-1})=f(g)\,e'[f(g)]^{-1}=f(g)[f(g)]^{-1}=e'.$$

$$\therefore \quad gkg^{-1}\in K.$$

Thus $g\in G,\ k\in K \Rightarrow gkg^{-1}\in K$.

$\therefore$ K is a normal subgroup of G.

Theorem 2. *Let f be a homomorphism of a group G into a group G' with kernel K. Let a be a given element of G such that $f(a)=a'\in G'$. Then the set of all those elements of G which have the image a' in G' is the coset Ka of K in G.* **(Meerut 1973)**

Proof. Let e, e' be the identities of G and G' respectively. Let $a\in G$ and $f(a)=a'\in G'$. Let $f^{-1}(a')=\{x\in G: f(x)=a'\}$. Then to prove that $f^{-1}(a')=Ka$.

Let $y\in Ka$. Then $y=ka$ for some $k\in K$.

We have $f(y)=f(ka)=f(k)f(a)$

$=e'f(a)$ $\qquad [\because k\in K\Rightarrow f(k)=e']$

$=f(a)=a'$.

$\therefore$ $y\in f^{-1}(a')$.

Thus $\quad y\in Ka\Rightarrow y\in f^{-1}(a')$.

$\therefore$ $Ka\subseteq f^{-1}(a')$. ...(1)

Now let z be any element of $f^{-1}(a')$. Then $f(z)=a'$.

We have $f(za^{-1})=f(z)f(a^{-1})=f(z)[f(a)]^{-1}=a'(a')^{-1}=e'$.

$\therefore$ $za^{-1}\in K$

$\Rightarrow (za^{-1})a\in Ka\Rightarrow z\in Ka$.

Thus $z\in f^{-1}(a')\Rightarrow z\in Ka$.

$\therefore$ $f^{-1}(a')\subseteq Ka$. ...(2)

From (1) and (2), we get $f^{-1}(a')=Ka$.

Theorem 3. *The necessary and sufficient condition for a homomorphism f of a group G into a group G' with kernel K to be an isomorphism of G into G' is that*

$K=\{e\}$. **(Kanpur 1980; Nagarjuna 80)**

Proof. Let f be a homomorphism of a group G into a group G'. Let e, e' be the identities of G, G' respectively. Let K be the kernel of f.

Suppose f is an isomorphism of G into G'. Then f is one-one. Let $a\in K$. Then

$f(a)=e'$ $\qquad$ [by def. of kernel]

$\Rightarrow f(a)=f(e)$ $\qquad [\because f(e)=e']$

$\Rightarrow a=e$ $\qquad [\because f$ is one-one$]$

Thus $a\in K\Rightarrow a=e$. In other words e is the only element of G which belongs to K. Therefore $K=\{e\}$.

Conversely suppose that $K=\{e\}$. Then to prove that f is an isomorphism of G into G' *i.e.*, to prove that f is 1-1.

If $a, b \in G$, then

$f(a)=f(b) \Rightarrow f(a)\,[f(b)]^{-1}=f(b)\,[f(b)]^{-1}$

$\Rightarrow f(a)\,f(b^{-1})=e'$ [$\because$ f is a homomorphism]

$\Rightarrow f(ab^{-1})=e'$ [$\because$ f is a homomorphism]

$\Rightarrow ab^{-1} \in K$ [by def. of kernel]

$\Rightarrow ab^{-1}=e$ [$\because$ $K=\{e\}$]

$\Rightarrow$ [illegible]$^{-1}b=eb \Rightarrow a=b$.

$\therefore$ f is one-one.

Hence f is an isomorphism of G into G'.

Theorem 4. *Suppose G is a group and N is a normal subgroup of G. Let f be a mapping from G to G/N defined by*

$$f(x)=Nx \;\forall\; x \in G.$$

Then f is a homomorphism of G onto G/N and kernel $f=N$.

(I.A.S. 1974; Sambalpur 77; Jabalpur 70; Mysore 73; Poona 73; Delhi 69)

Proof. Consider the mapping

$$f: G \to G/N \text{ such that } f(x)=Nx \;\forall\; x \in G.$$

Let Nx be any element of G/N. Then $x \in G$.

We have $f(x)=Nx$. Therefore the mapping f is onto G/N.

Let $a, b \in G$. Then

$f(ab)=Nab=(Na)(Nb)$ [$\because$ N is normal]

$=f(a)\,f(b)$.

$\therefore$ f is a homomorphism of G onto G/N.

Thus every quotient group of a group is a homomorphic image of the group. The mapping

$$f: G \to G/N \text{ such that } f(x)=Nx \;\forall\; x \in G$$

is called a **natural mapping** of G onto G/N.

Let K be the kernel of this homomorphism f. The identity of the quotient group G/N is the coset N. So $K=\{y \in G : f(y)=N\}$.

We shall prove that $K=N$.

Let $k \in K$. Then $f(k)=N$ *i.e.*, identity of G/N.

But by def. of f, we have $f(k)=Nk$.

Now $Nk=N \Rightarrow k \in N$.

Thus $k \in K \Rightarrow k \in N$. Therefore $K \subseteq N$.

Again let n be any element of N. Then $Nn=N$.

We have $f(n)=Nn=N$. Therefore $n \in K$.

Thus $n \in N \Rightarrow n \in K$.

$$\therefore \quad N \subseteq K.$$

Consequently $K=N$.

Theorem 5. Fundamental theorem on homomorphism of groups. *Every homomorphic image of a group G is isomorphic to some quotient group of G.* **(I.A.S. 1974; Allahabad 80; Andhra 77; Punjab 68; Kanpur 71; Jabalpur 79; Banaras 70; Madras 83; Meerut 79, 81, 83P, 84P)**

Proof. Let G' be the homomorphic image of a group G and f be the corresponding homomorphism. Then f is a homomorphism of G onto G'. Let K be the kernel of this homomorphism. Then K is a normal subgroup of G. We shall prove that

$$G/K \cong G'.$$

If $a \in G$, then $Ka \in G/K$ and $f(a) \in G'$. Consider the mapping $\phi : G/K \rightarrow G'$ such that $\phi\,(Ka)=f(a) \;\forall\; a \in G$.

First we shall show that the mapping ϕ is well-defined *i.e.*, if $a, b \in G$ and $Ka=Kb$, then $\phi\,(Ka)=\phi\,(Kb)$.

We have $Ka=Kb \Rightarrow ab^{-1} \in K$

$$\Rightarrow f\,(ab^{-1})=e' \qquad \text{(identity of } G')$$
$$\Rightarrow f\,(a) f\,(b^{-1})=e' \Rightarrow f\,(a)\,[f\,(b)]^{-1}=e'$$
$$\Rightarrow f\,(a)\,[\,f\,(b)]^{-1} f\,(b)=e' f(b)$$
$$\Rightarrow f\,(a)\,e'=f(b) \Rightarrow f(a)=f(b)$$
$$\Rightarrow \phi\,(Ka)=\phi\,(Kb).$$

$\therefore$ ϕ is well-defined.

ϕ is one-one. We have

$$\phi\,(Ka)=\phi\,(Kb) \Rightarrow f(a)=f(b)$$
$$\Rightarrow f(a)\,[f(b)]^{-1}=f(b)\,[f(b)]^{-1}$$
$$\Rightarrow f(a) f(b^{-1})=e' \Rightarrow f(ab^{-1})=e'$$
$$\Rightarrow ab^{-1} \in K. \qquad [\because \; K \text{ is kernel}]$$
$$\Rightarrow Ka=Kb.$$

$\therefore$ ϕ is one-one.

ϕ is **onto** G'. Let y be any element of G'. Then $y=f(a)$ for some $a \in G$ because f is onto G'. Now $Ka \in G/K$ and we have

$$\phi\,(Ka)=f(a)=y.$$

$\therefore$ ϕ is onto G'.

Finally we have $\phi\,[(Ka)\,(Kb\,]=\phi\,(Kab)=f\,(ab)$

$$=f(a) f(b)=\phi\,(Ka)\,\phi\,(Kb).$$

$\therefore$ ϕ is an isomorphism of G/K onto G'.

Hence $\qquad G/K \cong G'$.

Remark. The fundamental theorem on homomorphism of groups tells us how to find all possible homomorphic images of a given group G. Except for isomorphism these homomorphic images must be expressible in the form G/K where K is normal in G. Conversely for any normal subgroup N of G, G/N is a homomorphic image of G. Thus we have a one-to-one correspondence between the normal subgroups of G and the homomorphic images of G. Therefore to find all homomorphic images of G we should proceed as follows :

Find all normal subgroups N of G and construct all quotient groups G/N. The set of quotient groups so constructed gives us all homomorphic images of G (upto isomorphisms).

Solved Examples

Ex. 1. *Let f be a homomorphic mapping of a group G into a group G'. Let $f(G)$ be the homomorphic image of G in G'. Then $f(G)$ is a subgroup of G'.*

(Marathwada 1974; Jabalpur 70; Kerala 70; Meerut 74)

Solution. We have $f(G)=\{f(x) : x \in G\}$. Obviously $f(G) \subseteq G'$. Let a', b' be any two elements of $f(G)$. Then $f(a)=a'$, $f(b)=b'$ for some $a, b \in G$. We have

$$a'(b')^{-1}=f(a)[f(b)]^{-1}=f(a)f(b^{-1})$$
$$=f(ab^{-1}) \in f(G) \text{ since } ab^{-1} \in G.$$

Thus $a', b' \in f(G) \Rightarrow a'(b')^{-1} \in f(G)$.

$\therefore$ $f(G)$ is a subgroup of G'.

Ex 2. *Show that every homomorphic image of an abelian group is abelian and the converse is not true.*

Solution. Let G be an abelian group. Let f be a homomorphic mapping of G onto a group G'. Then G' is a homomorphic image of G.

Let a', b' be any two elements of G'. Then $f(a)=a'$, $f(b)=b'$ for some $a, b \in G$. We have

$$a'b'=f(a)f(b)=f(ab)$$
$$=f(ba) \qquad [\because G \text{ is abelian}]$$
$$=f(b)f(a)=b'a'.$$

$\therefore$ G' is abelian.

The converse is not true. P_3 is a non-abelian group. A_3 is a normal subgroup of P_3. The quotient group P_3/A_3 is a homomorphic image of P_3. Now P_3/A_3 is of order 2 and is abelian.

Ex. 3. *Show that a homomorphism from a simple group is either trivial or one-to-one.* **(Meerut 1978)**

Solution. Let G be a simple group and f be a homomorphism of G into another group G'. Then $\ker f$ is a normal subgroup of G. But the only normal subgroups of the simple group G are G itself and $\{e\}$. Therefore either $\ker f = G$ or $\ker f = \{e\}$. If $\ker f = G$, the f-image of each element of G is the identity of G' and so the homomorphism f is a trivial one. If $\ker f = \{e\}$, the homomorphism f is one-to-one. Hence the result.

Ex. 4. *If f is a homomorphism of G onto G' and g a homomorphism of G' onto G'', show that $g \circ f$ is a homomorphism of G onto G''. Also show that the kernel of f is a subgroup of that of $g \circ f$.*

(Allahabad 1979; Banaras 62, 64)

Solution. f is a mapping of G onto G' and g is a mapping of G' onto G''. Therefore $g \circ f$ is a mapping of G onto G'' and we have

$$(g \circ f)(x) = g[f(x)] \ \forall \ x \in G.$$

Let $a, b \in G$. Then

$$\begin{aligned}(g \circ f)(ab) &= g[f(ab)] \\ &= g[f(a)f(b)] && [\because f \text{ is a homomorphism}] \\ &= g[f(a)]\, g[f(b)] && [\because g \text{ is a homomorphism}] \\ &= [(g \circ f)(a)]\,[(g \circ f)(b)].\end{aligned}$$

$\therefore$ $g \circ f$ is a homomorphism of G onto G''.

Let K be the kernel of $g \circ f$. Then.

$K = \{y \in G : (g \circ f)(y) = e''$ where e'' is the identity of $G''\}$.

Let K' be the kernel of f. Then

$K' = \{z \in G : f(z) = e'$ where e' is the identity of $G'\}$.

Both K and K' are normal subgroups of G. In order to show that K' is a subgroup of K it is sufficient to prove that $K' \subseteq K$.

Let $k' \in K'$. Then $f(k') = e'$.

Also $(g \circ f)(k') = g[f(k')] = g(e') = e''$.

$\therefore$ $k' \in K$.

Thus $k' \in K' \Rightarrow k' \in K$.

$\therefore$ $K' \subseteq K$.

Ex. 5. *Let G be the multiplicative group of all $n \times n$ non-singular matrices with elements as real numbers and let G' be the multiplicative group of all non-zero real numbers. Show that the mapping*

$$f : G \to G' \text{ such that } f(A) = |A| \ \forall \ A \in G$$

is a homomorphism of G onto G'. What is the kernel of this homomorphism ?

Solution. Let A, B be any two $n \times n$ non-singular matrices with elements as real numbers. We have

$$f(AB) = |AB| = |A|\,|B| = f(A)\,f(B).$$

Also if x is any non-zero real number then there exists an $n \times n$ matrix $\in G$ whose determinant is equal to x.

$\therefore$ f is a homomorphism of G onto G'.

The identity of G' is 1. Therefore the kernel of f is the subgroup of G consisting of matrices with determinant equal to 1.

Ex. 6. *Show that the mapping* $f : \mathbf{C} \rightarrow \mathbf{R}$ *such that* $f(x+iy) = x$ *is a homomorphism of the additive group of complex numbers onto the additive group of real numbers and find the kernel of* f.

Solution. Let x_1+iy_1, x_2+iy_2 be any two elements of C. Then

$$\begin{aligned} f[(x_1+iy_1)+(x_2+iy_2)] &= f[(x_1+x_2)+i(y_1+y_2)] \\ &= x_1+x_2 \qquad [\text{by def. of } f] \\ &= f(x_1+iy_1)+f(x_2+iy_2). \end{aligned}$$

Also if x is any real number, then there exists a complex number $x+i0$ such that $f(x+i0) = x$.

$\therefore$ f is a homomorphism of C onto R.

The identity of R is the real number zero. Therefore the kernel of f consists of all complex numbers whose real part is zero.

Ex. 7. *If n be any given positive integer, show that the mapping*

$$f : \mathbf{C}_0 \rightarrow \mathbf{C}_0 \text{ defined by } f(z) = z^n$$

is an endomorphism of the multiplicative group of non-zero complex numbers. What is the kernel of this endomorphism ?

(Rajasthan 1967)

Solution. Let z_1, z_2 be any two elements of $\mathbf{C}_0$. Then

$$f(z_1) = z_1^n \text{ and } f(z_2) = z_2^n.$$

We have $\quad f(z_1 z_2) = (z_1 z_2)^n = z_1^n z_2^n = f(z_1) f(z_2).$

$\therefore$ f is an endomorphism of $\mathbf{C}_0$.

The identity of $\mathbf{C}_0$ is 1. The kernel of f consists of the n n^{th} roots of unity, *i.e.*, kernel of $f = \{e^{2r\pi i/n}, r = 0, 1, \ldots, n-1\}$.

We have, $f(e^{2r\pi i/n}) = (e^{2r\pi i/n})^n = e^{2r\pi i} = \cos 2r\pi + i \sin 2r\pi$

$$= 1 + 0i = 1, \text{ which is the identity of } \mathbf{C}_0.$$

Also $\quad f(z) = 1 \Rightarrow z^n = 1 \Rightarrow z = (1)^{1/n}$

$$\Rightarrow z \text{ is an } n^{th} \text{ root of unity.}$$

$\therefore$ the kernel of f consists of the n n^{th} roots of unity.

Ex. 8. *Let $\mathbf{C}_0$ and $\mathbf{R}_0$ be the multiplicative groups of non-zero complex numbers and of non-zero real numbers respectively. Then the mapping*

$f : \mathbf{C}_0 \to \mathbf{R}_0$ defined by $f(z) = |z| \ \forall \ z \in \mathbf{C}_0$

is a homomorphism of $\mathbf{C}_0$ into $\mathbf{R}_0$. What is the kernel of f?

(Raj. 1969; Bombay 70)

Solution. Let $z_1, z_2 \in \mathbf{C}_0$. Then $f(z_1) = |z_1|$, $f(z_2) = |z_2|$.

We have $f(z_1 z_2) = |z_1 z_2| = |z_1|\,|z_2| = f(z_1)\,f(z_2)$.

$\therefore$ f is a homomorphism of $\mathbf{C}_0$ into $\mathbf{R}_0$.

The identity of $\mathbf{R}_0$ is 1. The kernel of f is the multiplicative subgroup of all complex numbers whose modulus is 1.

Exercises

1. Let $f : G \to G'$ be a homomorphism of a group G into a group G'. Prove that $f(G)$ is a subgroup of G' and $f^{-1}(e')$ is a normal subgroup of G, where e' is the identity of G'. **(Kerala 1970)**

2. Prove that in a homomorphic mapping of a group G into a group G', unit element corresponds to unit element, inverses correspond to inverses and subgroups correspond to subgroups.

3. Let R be the additive group of real numbers and U, the multiplicative group of complex numbers of absolute value unity. Prove that the mapping $x \to e^{ix}$ is a homomorphism of R onto U. Find the kernel. **(Kerala 1970; Marathwada 74)**

Ans. Kernel $= \{x : x \in R$ and $x = 2n\pi$ where n is any integer$\}$.

4. Prove that any quotient group of G is a homomorphic image of G and conversely if G' is a homomorphic image of G then G' is isomorphic to a quotient group of G. **(Jabalpur 1970)**

5. G is a group and H is a subgroup of G. Show that the following statements are equivalent :

(*i*) H is a normal subgroup of G.

(*ii*) H is the kernel of a homomorphism of G.

(*iii*) Every left coset of H in G is a right coset of H in G.

(Marathwada 1972)

6. Prove that every normal subgroup of a group G gives rise to a homomorphism from G. Moreover, show that with every homomorphism from G we can associate a unique normal subgroup of G. **(Meerut 1972)**

7. If ϕ be a homomorphism of a group G onto the group $\bar{G}$ with kernel K, then prove that $G/K \cong \bar{G}$.

(Kanpur 1971; Vikram 76; Meerut 79, 81)

8. If ϕ is a homomorphism of a group G onto a group $\bar{G}$ with kernel K, then prove that the set of all inverse images of $\bar{g} \in \bar{G}$ under ϕ in G is given by Kx where x is any particular inverse image of $\bar{g}$ in G. **(Meerut 1973)**

9. Let H be a subgroup of a group G and S be the set of all right cosets of H in G. Prove that there is a homomorphism θ of G into $A(S)$ and the kernel of θ is the largest normal subgroup of G which is contained in H. **(Meerut 1973)**

10. Write down the elements of the symmetric group P_3 and determine the classes of conjugate elements. **(Bombay 1970)**

11. Show that any two conjugate classes of a group are either disjoint or identical. **(Delhi Hons. 1970)**

12. If N is a normal subgroup of G, having the prime index p, prove that G/N is cyclic.

13. Let $(G, +)$ be an abelian group. Let S be the set of all endomorphisms of G. For any $\sigma, \eta \in S$ define

$$\sigma+\eta : G \to G \text{ by}$$
$$(\sigma+\eta)(x)=\sigma(x)+\eta(x).$$

Show that $\sigma+\eta$ is also an endomorphism of G. Further show that S becomes an abelian group with respect to this addition composition. **(Meerut 1974 ; Andhra 75)**

14. Show that it is impossible to find a homomorphism of Z onto S_n $(n > 2)$. Here Z is the additive group of integers. **(Poona 1973)**

15. State and prove fundamental theorem of homomorphism for groups. Deduce that if a group G' is a homomorphic image of a finite group G, then the order of G' divides the order of G. **(Meerut 1976)**

16. How many homomorphisms are there from Z_4 onto Z_5 ? Here Z_n denotes the additive group of residue classes modulo n. **Ans.** No. **(Poona 1973)**

17. If the order of a group G is a power of a prime p, show that the centre of G has at least p elements. **(Madras 1983)**

§ 7. Automorphisms of a group.

Definition. **(Madras 1983 ; Meerut 78 ; Rajasthan 77)**

An isomorphic mapping of a group G onto itself is called an automorphism of G.

Thus $f : G \xrightarrow[\text{one-one}]{\text{onto}} G$ is an automorphism of G if

$$f(ab)=f(a)f(b) \ \forall \ a, b \in G.$$

Solved Examples

Ex. 1. *Show that the mapping*

$$f : \mathbf{I} \to \mathbf{I} \text{ such that } f(x)=-x \ \forall \ x \in \mathbf{I}$$

is an automorphism of the additive group of integers **I**.

Solution. Obviously the mapping f is one-one onto.

Let x_1, x_2 be any two elements of **I**. Then

$f(x_1+x_2)=-(x_1+x_2)=(-x_1)+(-x_2)=f(x_1)+f(x_2)$.

Hence f is an automorphism of I.

Ex. 2. *Show that $a \to a^{-1}$ is an automorphism of a group G iff G is abelian.*

[Nagarjuna 1978 ; Madras 78 ; Meerut 77, 78, 82, 83 (P), 84 (P)]

Solution. Let $f : G \to G$ be such that $f(x)=x^{-1}\ \forall\ x \in G$.

The function f is one-one because

$f(x)=f(y) \Rightarrow x^{-1}=y^{-1} \Rightarrow (x^{-1})^{-1}=(y^{-1})^{-1} \Rightarrow x=y$.

Also if $x \in G$, then $x^{-1} \in G$ and we have $f(x^{-1})=(x^{-1})^{-1}=x$.

∴ f is onto.

Now suppose G is abelian. Let a, b be any two elements of G.

Then $f(ab)=(ab)^{-1}$ [by def. of f]

$=b^{-1}a^{-1}=a^{-1}b^{-1}$ [∵ G is abelian]

$=f(a)f(b)$ [by def. of f]

∴ f is an automorphism of G.

Conversely suppose that f is an automorphism of G. Let $a, b \in G$.

We have $f(ab)=(ab)^{-1}$ [by def. of f]

$=b^{-1}a^{-1}=f(b)f(a)$ [by def of f]

$=f(ba)$. [∵ f is an automorphism]

Since f is one-one, therefore

$f(ab)=f(ba) \Rightarrow ab=ba \Rightarrow G$ is abelian.

Ex. 3. *Let G be a group, H a subgroup of G, f an automorphism of G. Let $f(H)=\{f(h) : h \in H\}$. Prove that $f(H)$ is a subgroup of G.*

Solution. Let a, b be any two elements of $f(H)$ Then

$a=f(h_1)$ and $b=f(h_2)$ where $h_1, h_2 \in H$.

Now $h_1, h_2 \in H \Rightarrow h_1h_2^{-1} \in H$ [H is a subgroup]

$\Rightarrow f(h_1h_2^{-1}) \in f(H)$

$\Rightarrow f(h_1)f(h_2^{-1}) \in f(H)$

[f is an automorphism]

$\Rightarrow f(h_1)[f(h_2)]^{-1} \in f(H) \Rightarrow ab^{-1} \in f(H)$.

∴ $f(H)$ is a subgroup of G.

Note. Some authors use the symbol hf in place of $f(h)$ to denote the image of an element.

Ex. 4. *Let G be a group, f an automorphism of G, N a normal subgroup of G. Prove that $f(N)$ is a normal subgroup of G.*

Solution. First show as in Ex. 3 that $f(N)$ is a subgroup of G.

Now to show that $f(N)$ is a normal subgroup of G.

Let $x \in G$ and $k \in f(N)$. Then $x=f(y)$ where $y \in G$ because f is a function of G onto G. Also $k=f(n)$ where $n \in N$.

We have

$$xkx^{-1}=f(y)f(n)[f(y)]^{-1}=f(y)f(n)f(y^{-1})=f(yny^{-1}).$$

Since N is normal in G, therefore $yny^{-1} \in N$. Consequently $f(yny^{-1}) \in f(N)$. Thus $xkx^{-1} \in f(N)$.

$\therefore$ $f(N)$ is a normal subgroup of G.

Group of automorphisms of a group.

Theorem. *The set of all automorphisms of a group forms a group with respect to composite of functions as the composition.*

(Gujrat 1970, 71 ; Banaras 61 ; Madras 78 ; Raj. 77)

Proof. Let $A(G)$ be the collection of all automorphisms of a group G. Then $A(G)=\{f : f \text{ is an automorphism of } G\}$.

We shall prove that $A(G)$ is a group with respect to composite of functions as composition.

Closure property Let $f, g \in A(G)$. Then f, g are one-one mappings of G onto itself. Therefore gf is also a one-one mapping of G onto itself. If a, b be any two elements of G, we have

$$(gf)(ab)=g[f(ab)]=g[f(a)f(b)]$$
$$=g[f(a)]\,g[f(b)]=[(gf)(a)][(gf)(b)].$$

$\therefore$ gf is also an automorphism of G. Thus $A(G)$ is closed with respect to composite composition.

Associativity. We know that composite of arbitrary mappings is associative. Therefore composite of automorphisms is also associative.

Existence of Identity. The identity function i on G is also an automorphism of G. Obviously i is one-one onto and if $a, b \in G$, then $i(ab)=ab=i(a)\,i(b)$. Thus $i \in A(G)$ and if $f \in A(G)$, we have $if=f=fi$.

Existence of Inverse. Let $f \in A(G)$. Since f is a one-one mapping of G onto itself, therefore f^{-1} exists and is also a one-one mapping of G onto itself. We shall show that f^{-1} is also an automorphism of G. Let $a, b \in G$. Then there exist $a', b' \in G$ such that

$$f^{-1}(a)=a' \Leftrightarrow f(a')=a$$
$$f^{-1}(b)=b' \Leftrightarrow f(b')=b.$$

We have $\quad f^{-1}(ab)=f^{-1}[f(a')f(b')]$

$$=f^{-1}[f(a'b')]=a'b'=f^{-1}(a)\,f^{-1}(b).$$

$\therefore$ f^{-1} is an automorphism of G and thus

$$f \in A(G) \Rightarrow f^{-1} \in A(G).$$

Therefore each element of $A(G)$ possesses inverse.

Therefore $A(G)$ is a group with respect to composite composition.

§ 8. Inner Automorphisms. We shall now study a special type of automorphisms known as **inner automorphisms**. First we shall prove a preliminary theorem.

Theorem 1. *Let a be a fixed element of a group G. Then the mapping* $f_a : G \to G$ *defined by* $f_a(x) = a^{-1}xa \; \forall \; x \in G$ *is an automorphism of G.* **(Gujrat 1971)**

Proof. **The mapping f_a is one-one.** Let x, y be any two elements of G. Then

$f_a(x) = f_a(y) \Rightarrow a^{-1}xa = a^{-1}ya \Rightarrow x = y$, by cancellation laws in G. Therefore the mapping f_a is one-one.

The mapping f_a is also onto G. If y is any element of G, then $aya^{-1} \in G$ and we have $f_a(aya^{-1}) = a^{-1}(aya^{-1})\, a = y$.

$\therefore$ f_a is onto G.

Finally if $x, y \in G$ then $f_a(xy) = a^{-1}(xy)\, a = (a^{-1}xa)(a^{-1}ya) = f_a(x)\, f_a(y)$. Hence f_a is an automorphism of G.

Inner Automorphism. Definition.

If G is a group, the mapping

$$f_a : G \to G \text{ defined by } f_a(x) = a^{-1}xa \; \forall \; x \in G$$

is an automorphism of G known as inner automorphism.

(Delhi 1964; Nagarjuna 78; Aligarh 65)

Also an automorphism which is not inner is called an **outer automorphism.**

Theorem 2. *For an abelian group the only inner automorphism is the identity mapping whereas for non-abelian groups there exist non-trivial automorphisms.* **(Raj. M. Sc. 1966)**

Proof. Suppose G is an abelian group and f_a is an inner automorphism of G. If $x \in G$, we have

$$f_a(x) = a^{-1}xa = a^{-1}ax \qquad [\because G \text{ is abelian}]$$
$$= ex = x.$$

Thus $f_a(x) = x \; \forall \; x \in G$.

$\therefore$ f_a is the identity mapping of G.

Let now G be non-abelian. Then there exist at least two elements say $a, b \in G$ such that

$$ba \neq ab \Rightarrow a^{-1}ba \neq b \Rightarrow f_a(b) \neq b.$$

Hence f_a is not the identity mapping of G. Thus for non-abelian groups there always exist non-trivial inner automorphisms.

Theorem 3. *The set $I(G)$ of all inner automorphisms of a group G is a normal subgroup of the group of its automorphisms isomorphic to the quotient group G/Z of G where Z is the centre of G.*

(I. A. S. 1970; Delhi 70; Vikram 76; Nagarjuna 78; Madras 83; Gujrat 71; Dibrugarh 78; Meerut 74, 78, 79)

Proof. Let $A(G)$ denote the group of all automorphisms of G. Then $I(G) \subseteq A(G)$.

Let $a, b \in G$. We shall first prove the following two results :

(i) $f_{a^{-1}} = f_a^{-1}$ *i.e.*, the inner automorphism $f_{a^{-1}}$ is the inverse function of the inner automorphism f_a.

(ii) $f_a f_b = f_{bc}$.

Proof of (i). If $x \in G$, then we have

$$(f_a f_{a^{-1}})(x) = f_a[f_{a^{-1}}(x)] = f_a[(a^{-1})^{-1} xa^{-1}] = f_a[axa^{-1}]$$

$$= a^{-1}(axa^{-1})a = x.$$

$\therefore$ $f_a f_{a^{-1}}$ is the identity function on G.

$$\therefore \quad f_{a^{-1}} = (f_a)^{-1}.$$

Proof of (ii). If $x \in G$, then we have

$$(f_a f_b)(x) = f_a[f_b(x)] = f_a(b^{-1}xb) = a^{-1}(b^{-1}xb)a = (a^{-1}b^{-1})x(ba)$$

$$= (ba)^{-1}x(ba) = f_{ba}(x).$$

$$\therefore \quad f_a f_b = f_{ba}.$$

Now we shall prove that $I(G)$ is a subgroup of $A(G)$. Let f_a, f_b be any two elements of $I(G)$. Then

$$f_a(f_b)^{-1} = f_a f_{b^{-1}} = f_{b^{-1}a} \in I(G) \text{ since } b^{-1}a \in G.$$

Thus $f_a, f_b \in I(G) \Rightarrow f_a(f_b)^{-1} \in I(G)$.

$\therefore$ $I(G)$ is a subgroup of $A(G)$.

Now we shall prove that $I(G)$ is a normal subgroup of $A(G)$.

Let $f \in A(G)$ and $f_a \in I(G)$. If $x \in G$, then we have

$$(f f_a f^{-1})(x) = (f f_a)[f^{-1}(x)] = f[f_a(f^{-1}(x))]$$

$$= f[a^{-1} f^{-1}(x)\, a]$$

$$= f(a^{-1}) f[f^{-1}(x)] f(a) \qquad [\because f \text{ is composition preserving}]$$

$$= f(a^{-1})\, x f(a) \qquad [\because f[f^{-1}(x)] = x]$$

$$= [f(a)]^{-1} x f(a)$$

$$= c^{-1}xc \text{ where } f(a) = c \in G$$

$$= f_c(x).$$

$\therefore$ $f f_a f^{-1} = f_c \in I(G)$ since $c \in G$.

$\therefore$ $I(G)$ is a normal subgroup of $A(G)$.

Now we shall show that $I(G)$ is isomorphic to G/Z. For this we shall show that $I(G)$ is a homomorphic image of G and Z is the kernel of the corresponding homomorphism.

Then by the fundamental theorem on homomorphism of groups we shall have $G/Z \cong I(G)$.

Consider the mapping $\phi : G \to I(G)$ defined by

$$\phi(a) = f_{a^{-1}} \ \forall \ a \in G.$$

Obviously ϕ is onto $I(G)$ because $f_a \in I(G) \Rightarrow a \in G$ and this implies $a^{-1} \in G$.

Now $$\phi(a^{-1}) = f_{(a^{-1})^{-1}} = f_a.$$

$\therefore$ ϕ is onto $I(G)$.

Now to prove that $\phi(ab) = \phi(a)\,\phi(b) \ \forall \ a, b \in G$.

We have $\phi(ab) = f_{(ab)^{-1}} = f_{b^{-1}a^{-1}} = f_{a^{-1}}\, f_{b^{-1}} = \phi(a)\phi(b)$.

Now to show that Z is the kernel of ϕ.

The identity function i on G is the identity of the group $I(G)$.

Let K be the kernel of ϕ.

Then we have $z \in K \Leftrightarrow \phi(z) = i \Leftrightarrow f_{z^{-1}} = i \Leftrightarrow f_{z^{-1}}(x) = i(x)$
$\forall \ x \in G \Leftrightarrow (z^{-1})^{-1} x z^{-1} = x \ \forall \ x \in G \Leftrightarrow z x z^{-1} = x \ \forall \ x \in G$
$\Leftrightarrow zx = xz \ \forall \ x \in G \Leftrightarrow z \in Z$.

$\therefore$ $K = Z$.

Hence the theorem.

§ 9. Group of automorphisms of a cyclic group.

(Bombay 1970; Madras 78)

Suppose $G = \{a\}$ is a cyclic group generated by a. An automorphism f of G is completely defined by a relation of the form

$$f(a) = a^m, \qquad \ldots(1)$$

where m is some suitable integer.

For if k is any integer, then for f to be an automorphism of G, we have

$$f(a^k) = [f(a)]^k - (a^m)^k. \qquad \ldots(2)$$

The relation (2) gives the f-image of each element of G and the mapping f is thus completely defined.

Now let b be any element of G. Since f is a mapping of G onto itself, therefore there must exist an element $a^k \in G$ such that

$$b = f(a^k) = (a^m)^k. \qquad \ldots(3)$$

From (3) we conclude that for the mapping f defined in (1) to be an automorphism of G, a^m must be a generator of G.

Now if G is infinite, the only generators of G are a and a^{-1}. So in this case the only automorphisms of G are, (i) the identity mapping I for which

$$I(a) = a \Rightarrow I(a^k) = a^k \ \forall \ k \in \mathbf{I}$$

and, (ii) the mapping f defined by

$$f(a) = a^{-1}.$$

Therefore the group of automorphisms of an infinite cyclic group is of order 2. **(Meerut 1979)**

On the other hand if G is of finite order n, then a^m is a generator of G if and only if m is prime to n and less than n. We shall show that for each such m, the mapping f defined in (1) is an automorphism of G.

f **is one-one.** Let a^{k_1} and a^{k_2} be any two elements of G where $1 \leqslant k_1 \leqslant n$, $1 \leqslant k_2 \leqslant n$.

Then $f(a^{k_1})=f(a^{k_2}) \Rightarrow [f(a)]^{k_1}=[f(a)]^{k_2} \Rightarrow a^{mk_1}=a^{mk_2}$
$\Rightarrow a^{m(k_1-k_2)}=e \Rightarrow n \mid m\,(k_1-k_2)$.

But m is prime to n and $0 \leqslant (k_1-k_2) < n$. Therefore
$n \mid m\,(k_1-k_2) \Rightarrow k_1-k_2=0 \Rightarrow k_1=k_2 \Rightarrow a^{k_1}=a^{k_2}$. Thus

$$f(a^{k_1})=f(a^{k_2}) \Rightarrow a^{k_1}=a^{k_2}.$$

Therefore the mapping f is one-one.

f **is onto.** Since G is finite and f is one-one, therefore f must be onto G.

Finally if a^{k_1}, a^{k_2} are any two elements of G, then

$$\begin{aligned} f(a^{k_1}\,a^{k_2}) &= f(a^{k_1+k_2})=[f(a)]^{k_1+k_2}=a^{m\,(k_1+k_2)} \\ &= a^{mk_1}\,a^{mk_2}=(a^m)^{k_1}\,(a^m)^{k_2}=[f(a)]^{k_1}\,[f(a)]^{k_2} \\ &= f(a^{k_1})\,f(a^{k_2}). \end{aligned}$$

Therefore the mapping f denoted in (1) is an automorphism for each positive integer m less than n and prime to n.

Hence the group of automorphisms of a finite cyclic group of order n is of order $\phi(n)$ where $\phi(n)$ denotes the number of integers less than n and prime to n. **(Poona 1970)**

In the end we shall show that *the group of automorphisms of a cyclic group is abelian.*

Let f_{m_1}, f_{m_2} be two automorphisms defined by

$$f_{m_1}(a)=a^{m_1},\ f_{m_2}(a)=a^{m_2}.$$

$$\begin{aligned} \text{Then } (f_{m_1} \circ f_{m_2})(a) &= f_{m_1}[f_{m_2}(a)]=f_{m_1}(a^{m_2}) \\ &= [f_{m_1}(a)]^{m_2}=(a^{m_1})^{m_2}=a^{m_1m_2} \\ &= (a^{m_2})^{m_1}=[f_{m_2}(a)]^{m_1}=f_{m_2}(a^{m_1}) \\ &= f_{m_2}[f_{m_1}(a)]=(f_{m_2} \circ f_{m_1})(a). \end{aligned}$$

Now two automorphisms of a cyclic group are equal if the image of a generator of the group under each of them is the same.

Hence $f_{m_1} \circ f_{m_2} = f_{m_2} \circ f_{m_1}$. Therefore the group of automorphisms of a cyclic group is abelian.

Exercises

1. Let G be a finite abelian additive group and n be a positive integer relatively prime to $o(G)$. Prove that the mapping $\sigma : G \to G$ given by $\sigma(x) = nx$ is an automorphism of G. (Meerut 1974)
2. Verify the following statement for being true or false :
   If $G=(a)$ is a cyclic group of order 10 then the mapping $\sigma : G \to G$ such that $\sigma(a^k)=a^{2k}$ for all k, is an automorphism of G. (Meerut 1976)
   Ans. False.
3. Give an example of a group in which (i) the inner automorphisms corresponding to any two elements are the same, (ii) the inner automorphisms corresponding to no two elements are the same. (I.A.S. 1975)
   Ans. (i) Every abelian group, (ii) the symmetric group P_3.
4. Show that the group of all automorphisms of a cyclic group G of order r is isomorphic to the group of integers less than and relatively prime to r under multiplication modulo r. (I.A.S. 1970)

§ 10. More results on group homomorphism.

Theorem 1. *Let G be a group and H a normal subgroup of G. If K is a normal subgroup of G containing H i.e., $H \subseteq K$, then the quotient group K/H is a normal subgroup of the quotient group G/H. Conversely, if K/H is a normal subgroup of G/H, then K is a normal subgroup of G containing H.*

Proof. It is given that H is a normal subgroup of G and $H \subseteq K$ where K itself is a normal subgroup of G. Therefore H is also a normal subgroup of K and consequently K/H is a quotient group.

If the coset Ha is an element of K/H, then a is an element of K. Now $a \in K \Rightarrow a \in G$. Therefore Ha is also an element of G/H. Thus the quotient group K/H is a subset of the quotient group G/H. Therefore K/H, is a subgroup of G/H.

We shall now show that K/H is normal in G/H. Let Hg be any element of G/H and Hk be any element of K/H. Then $g \in G$ and $k \in K$. We have

$(Hg)(Hk)(Hg)^{-1} = (Hg)(Hk)(Hg^{-1})$ $[\because (Hg)^{-1} = Hg^{-1}]$

$= Hgkg^{-1}$ $[\because H$ is normal $\Rightarrow (Ha)(Hb) = Hab]$

Since K is a normal subgroup of G, therefore $gkg^{-1} \in K$. Consequently $Hgkg^{-1} \in K/H$. Therefore K/H is a normal subgroup of G/H.

Conversely, let x be any element of G and let k be any element of K. In order to prove that K is normal in G we must show that xkx^{-1} is in K.

We have $Hx \in G/H$ and $Hk \in K/H$. Since K/H is given to be a normal subgroup of G/H, therefore

$$(Hx)\,(Hk)\,(Hx)^{-1} \in K/H$$

$$\Rightarrow Hxkx^{-1} \in K/H \qquad [\because H \text{ is normal in } G]$$

$$\Rightarrow xkx^{-1} \in K.$$

$\therefore$ K is normal in G. Also K/H is a quotient group implies that H is a normal subgroup of K. Therefore K is a normal subgroup of G and $H \subseteq K$.

Theorem 2. *If H be a normal subgroup of a group G and K a normal subgroup of G containing H, then $G/K \cong (G/H)/(K/H)$.*

(I.A.S. 1971)

Proof. Since H is a normal subgroup of G and K is a normal subgroup of G containing H, therefore by theorem 1, the quotient group K/H is a normal subgroup of the quotient group G/H. Hence $(G/H)/(K/H)$ is a quotient group.

Now consider the mapping $\phi : G/H \rightarrow G/K$ defined by

$$\phi\,(Hx) = Kx,\ x \in G.$$

We shall first show that ϕ is well-defined.

Let $\qquad Hx = Hy$ where $x, y \in G$.

Then
$$xy^{-1} \in H$$

$$\Rightarrow xy^{-1} \in K \qquad [\because H \subseteq K]$$

$$\Rightarrow Kx = Ky \Rightarrow \phi\,(Hx) = \phi\,(Hy).$$

$\therefore$ ϕ is well-defined.

Now we shall show that ϕ is a homomorphism of G/H onto G/K. If $x, y \in G$ then

$$\phi\,[(Hx)\,(Hy)] = \phi\,(Hxy) = Kxy = (Kx)\,(Ky) = \phi\,(Hx)\,\phi\,(Hy).$$

Also ϕ is obviously onto G/K because $Kx \in G/K$ implies that there exists $Hx \in G/H$ such that $\phi\,(Hx) = Kx$.

$\therefore$ ϕ is a homomorphism of G/H onto G/K.

Let us now find the kernel of ϕ. We claim that the kernel of $\phi = K/H$ which is obviously a subset of G/H. The proof is as follows :

The identity element of the group G/K is K. If $Hx \in G/H$, then $\qquad Hx \in$ the kernel of ϕ

$\Leftrightarrow \phi(Hx)=K \Leftrightarrow Kx=K \Leftrightarrow x \in K \Leftrightarrow Hx \in K/H.$

Therefore the kernel of $\phi=K/H$.

Now ϕ is a homomorphism of G/H onto G/K with kernel K/H. Therefore by the fundamental theorem on homomorphism of groups, we have $G/K \cong (G/H)/(K/H)$.

Note. This theorem is known as the **Second law of isomorphism.**

Theorem 3. *Let G be a group and let H be any subgroup of G. If N is any normal subgroup of G, then*

$$HN/N \cong H/(H \cap N).$$

(Vikram 1976; Kanpur 80; Meerut 74)

Proof. If H is a subgroup of G and N is a normal subgroup of G, then we know that $H \cap N$ is a normal subgroup of H and consequently $H/(H \cap N)$ is a quotient group. [See Ex. 5, Page 192].

Also HN is a subgroup of G and $N \subseteq HN$. Since N is normal in G, therefore N is also a normal subgroup of HN. Hence HN/N is a quotient group. Now consider the mapping

$$\phi : H \to HN/N \text{ defined by } \phi(x)=Nx,\ x \in H.$$

The mapping ϕ is well-defined since $H \subseteq HN$ and therefore $x \in H \Rightarrow x \in HN$. Thus $Nx \in HN/N$.

Now we shall show that ϕ is a homomorphism of H onto HN/N with kernel $H \cap N$.

ϕ is onto. Let Nx be any member of HN/N.

Then $x \in HN$. Therefore $x=hn$ for some $h \in H$ and $n \in N$.

Since N is normal in G, therefore $HN=NH$.

Thus there exist $n' \in N$, $h' \in H$ such that $hn=n'h'$.

We have $\phi(h')=Nh'$

$$=(Nn')h' \qquad [\because n' \in N \Rightarrow Nn'=N]$$

$$=N(n'h')=N(hn)=Nx.$$

Thus $Nx \in HN/N \Rightarrow \exists\, h' \in H$ such that $\phi(h')=Nx$.

$\therefore$ ϕ is onto HN/N.

Again $\phi(xy)=Nxy=(Nx)(Ny)=\phi(x)\,\phi(y)$

$\therefore$ **ϕ is a homomorphism of H onto HN/N.**

Now to show that the kernel of $\phi=H \cap N$ which is obviously a subset of H.

The identity element of the group HN/N is N. If $h \in H$, then $h \in$ the kernel of $\phi \Rightarrow \phi(h)=N \Rightarrow Nh=N$

$$\Rightarrow h \in N$$

$$\Rightarrow h \in H \cap N \qquad [\because h \in H]$$

Again if $h \in H \cap N$, then $h \in H$ and $h \in N$.

Now $\phi(h)=Nh$

$=N$ $[\because h\in N]$

$\therefore$ $h\in H\cap N\Rightarrow h\in$ the kernel of ϕ.

Thus if $h\in H$, then $h\in$ the kernel of ϕ iff $h\in H\cap N$. Therefore the kernel of $\phi=H\cap N$.

Now by the fundamental theorem on homomorphism of groups, we have $HN/N\cong H/(H\cap N)$.

Note. This theorem is known as the **Third law of isomorphism.**

Theorem 4. *Let f be a homomorphism of a group G onto a group G' with kernel H. For each subgroup K' of G', define K by $K=\{x\in G: f(x)\in K'\}$. Then*

(*i*) *K is a subgroup of G containing H and $K'\cong K/H$.*

(*ii*) *K' is normal in G' iff K is normal in G.* **(Gujrat 1970)**

(*iii*) *If K' is normal in G', then $G/K\cong G'/K'$.* **(Meerut 1970)**

(*iv*) *$K'\leftrightarrow K$ is a one-one correspondence between the set of all subgroups of G' and the set of all subgroups of G which contain H.*

Proof. (*i*) Let a, b be any two elements of K.

Then $f(a), f(b)\in K'$.

Now K' is a subgroup of G'. Therefore

$$\begin{aligned}f(a), f(b)\in K' &\Rightarrow f(a)\,[f(b)]^{-1}\in K'\\ &\Rightarrow f(a)\,f(b^{-1})\in K' \quad [\because f \text{ is a homomorphism}]\\ &\Rightarrow f(ab^{-1})\in K'\\ &\Rightarrow ab^{-1}\in K. \quad [\text{by def. of } K]\end{aligned}$$

Thus $a, b\in K\Rightarrow ab^{-1}\in K$.

Therefore K is a subgroup of G.

If $h\in$ kernel H of f, then $f(h)=e'$ (identity of G')

$\in K'$.

$\therefore$ $h\in K$ and consequently $H\subseteq K$

Therefore K is a subgroup of G containing H.

The kernel H of f is a normal subgroup of G Therefore H is also a normal subgroup of K. The homomorphism f of G onto G', when considered only on the elements of K, induces a 'homomorphism of K onto K' with kernel H because $H\subseteq K$. Therefore by the fundamental theorem on homomorphism of groups, we have

$$K'\cong K/H.$$

(ii) First to show that if K' is normal in G', then K is normal in G. Let $x\in G$ and $k\in K$. Then

$$f(xkx^{-1})=f(x)\,f(k)\,f(x^{-1})=f(x)\,f(k)\,[f(x)]^{-1}.$$

Since $f(x)\in G'$ and $f(k)\in K'$ and K' is normal in G', therefore $f(x)\,f(k)\,[f(x)]^{-1}\in K'$.

Thus $f(xkx^{-1}) \in K' \Rightarrow xkx^{-1} \in K \Rightarrow K$ is normal in G.

Now to show that if K is normal in G, then K' is normal in G'. Let $y \in G'$ and $k' \in K'$. Then

$y=f(x)$ for some $x \in G$ and $k'=f(k)$ for some $k \in K$.

We have $yk'y^{-1}=f(x)f(k)[f(x)]^{-1}=f(x)f(k)f(x^{-1})=f(xkx^{-1})$.

Since K is normal in G, therefore $xkx^{-1} \in K$. Consequently $(xkx^{-1}) \in K'$. Thus $yk'y^{-1} \in K'$ $\forall$ $y \in G'$ and $\forall$ $k' \in K'$.

$\therefore$ K' is a normal subgroup of G'.

(iii) First prove as in parts (i) and (ii) that if K' is normal in G', then K is normal in G. Now define a function

$$\phi : G \rightarrow G'/K' \text{ such that } \phi(x)=K'f(x) \ \forall \ x \in G.$$

We shall show that ϕ is a homomorphism of G onto G'/K' with kernel K.

ϕ is onto. Let $K'x' \in G'/K'$. Then $x' \in G'$. Therefore $\exists$ $x \in G$ such that $f(x)=x'$.

Now $\phi(x)=K'f(x)=K'x'$. Therefore ϕ is onto G'/K'.

Again let $a, b \in G$. Then

$\phi(ab)=K'f(ab)=K'f(a)f(b)=[K'f(a)][K'f(b)]=\phi(a)\phi(b)$.

$\therefore$ ϕ is a homomorphism of G onto G'/K'.

Now to show that the kernel of $\phi=K$ The identity element of the group G'/K' is K'. If $g \in G$, then $g \in$ the kernel of $\phi \Leftrightarrow \phi(g) = K' \Leftrightarrow K'f(g)=K' \Leftrightarrow f(g) \in K' \Leftrightarrow g \in K$.

Therefore the kernel of $\phi=K$. Hence by the fundamental theorem on homomorphism of groups, we have $G/K \cong G'/K'$.

(iv) Let T be a subgroup of G containing H.

Let $T'=\{f(t) \in G' : t \in T\}$.

Then T' is a subgroup of G' as shown below :

Let $f(t_1), f(t_2) \in T'$. Then $t_1, t_2 \in T$

Now $t_1, t_2 \in T \Rightarrow t_1 t_2^{-1} \in T$ [$\because$ T is a subgroup]

$\Rightarrow f(t_1 t_2^{-1}) \in T'$ [by def. of T']

$\Rightarrow f(t_1) f(t_2^{-1}) \in T'$ [$\because$ f is a homomorphism]

$\Rightarrow f(t_1) [f(t_2)]^{-1} \in T'$.

$\therefore$ T' is a subgroup of G'.

Now let $L=\{l \in G : f(l) \in T'\}$. Then L is a subgroup of G containing H as shown in part (i). Our claim is that $L=T$.

Obviously $T \subseteq L$ since $t \in T \Rightarrow f(t) \in T' \Rightarrow t \in L$.

Now to show that $L \subseteq T$.

Let $l \in L$. Then $f(l) \in T'$

$\Rightarrow l \in T.$ [by def. of T]

$\therefore \quad L \subseteq T.$

Hence $\quad L = T.$

Thus there exists a one-one correspondence between the set of all subgroups of G' and the set of all subgroups of G which contain H.

Note. Some authors use the notation $f^{-1}(K')$ or $K'f^{-1}$ to denote $K = \{x \in G : f(x) \in K'\}$.

Ex. *Let ϕ be a homomorphism of a group G onto a group $\bar{G}$ with kernel K. Let $\bar{N}$ be a normal subgroup of $\bar{G}$ and let*

$$N = \{x \in G : \phi(x) \in \bar{N}\}.$$

Prove that N is a normal subgroup of G and

$$G/N \cong \bar{G}/\bar{N} \cong G/K/N/K. \qquad \textbf{(Meerut 1974)}$$

§ 11. Maximal subgroups.

Definition. *A normal subgroup H of a group G is said to be maximal if there exists no proper normal subgroup K of G which properly contains H.* **(Andhra 1977)**

Thus a normal subgroup H of a group G is maximal if and only if there exists no normal subgroup K of G such that

$$H \subset K \subset G$$

where the symbol $\subset$ stands for **proper inclusion.**

Theorem. *A normal subgroup H of G is maximal if and only if the quotient group G/H is simple.* **(Andhra 1977; Nagarjuna 78)**

Proof. Suppose H is maximal and G/H is not simple *i.e.*, G/H possesses proper normal subgroups. It should be noted that a group is said to be simple if it possesses no proper normal subgroups. Let K/H be a proper normal subgroup of G/H. Then by theorem 1 of § 10, K will be a normal subgroup of G containing H. Since K/H is a proper subgroup of G/H, therefore $H \subset K \subset G$. Thus K is a normal subgroup of G and $H \subset K \subset G$. Therefore H is not maximal. This contradicts the hypothesis that H is maximal in G. Hence G/H must be simple.

Conversely, let G/H be simple and let H be not maximal. Since H is not maximal, therefore there exists a normal subgroup K of G such that $H \subset K \subset G$.

Then by theorem 1 of § 10, K/H is a normal subgroup of G/H. Since $H \subset K \subset G$, therefore K/H is a proper normal subgroup of G/H *i.e.*, neither K/H is equal to the entire group G/H nor K/H is equal to the identity subgroup H/H. Consequently G/H is not simple. This contradicts the hypothesis that G/H is simple. Hence H must be maximal in G.

§ 12. Composition series of a group and the Jordan-Holder theorem.

(I.A.S. 1971, 74 ; Andhra 77)

Definition. *Let G be a group. Then a finite sequence of its subgroups*

$$G=H_1, H_2, H_3, \dots, H_n=\{e\} \qquad \dots(1)$$

is called a **composition series** *for G if each H_i except H_1 is a maximal normal subgroup of H_{i-1}.*

The quotient groups $G/H_2, H_2/H_3, \dots, H_{n-1}/H_n$ which are necessarily simple are then called **composition factor groups** or **composition quotient groups** of the composition series (1).

Example 1. **(Andhra 1977)**

Let $G=P_3=\{I, (12), (23), (31), (123), (132)\}$ and let $H_2=\{I, (123), (132)\}$. Then $G, H_2, \{I\}$ is a composition series for G.

Obviously H_2 is a maximal normal subgroup of G and $\{I\}$ is a maximal normal subgroup of H_2.

Example 2. Let G be a cyclic group of order 6 generated by a *i.e.*, let $G=\{a, a^2, a^3, a^4, a^5, a^6=e\}$. Then

$$G, H_2=\{e, a^3\}, \{e\} \text{ and } G, N_2=\{e, a^2, a^4\}, \{e\}$$

are two different composition series for G.

Example 3. Let G be a cyclic group $\{a\}$ of order 12 generated by a. Then

$$\{a\}, \{a^2\}, \{a^4\}, \{e\} \text{ and } \{a\}, \{a^3\}, \{a^6\}, \{e\}$$

are two different composition series for G.

Theorem 1. *There exists at least one composition series for every finite group G.* **(I.A.S. 1974 ; Lucknow 70)**

Proof. (i) If G is simple, then $G, \{e\}$ is a composition series for G.

(ii) Suppose G is not simple. Then there exists a proper normal subgroup H of G. If H is maximal in G and $\{e\}$ is maximal in H, then $G, H, \{e\}$ is a composition series.

Suppose H is not maximal in G but $\{e\}$ is maximal in H. Then there exists a normal subgroup K of G such that $G \supset K \supset H$. If K is maximal in G and H is maximal in K then $G, K, H, \{e\}$ is a composition series.

Now suppose that H is maximal in G but $\{e\}$ is not maximal in H. Then there exists a normal subgroup J of H such that

$$H \supset J \supset \{e\}.$$

If $\{e\}$ is maximal in J and J is maximal in H, then $G, H, J, \{e\}$ is a composition series.

Next suppose that H is not maximal in G and $\{e\}$ is not maximal in H. Then there exists a normal subgroup L of G such that $G \supset L \supset H$. Also there exists a normal subgroup N of H such that $H \supset N \supset \{e\}$. Thus $G \supset L \supset H \supset N \supset \{e\}$. If L is maximal in G, H is maximal in L, N is maximal in H and $\{e\}$ is maximal in N, then $G, L, H, N, \{e\}$ is a composition series.

Since G is finite, there are only a finite number of subgroups and ultimately we must reach a composition series.

Theorem 2. Jordan-Holder Theorem. *Let G be a finite group with two composition series*

$$G, H_1, H_2, \ldots, H_n = \{e\} \qquad \ldots(1)$$

and

$$G, K_1, K_2, \ldots, K_m = \{e\}. \qquad \ldots(2)$$

Then $n=m$ and the two corresponding series of composition quotient groups, viz.,

$$G/H_1, H_1/H_2, \ldots, H_{n-1}/H_n$$

and

$$G/K_1, K_1/K_2, \ldots, K_{m-1}/K_m$$

are abstractly identical i.e., they can be put into 1—1 correspondence such that the corresponding quotient groups are isomorphic.

(I.A.S. 1975 ; Banaras 74 ; Kurushetra 68 ; Nagpur 70)

Proof. We shall prove the theorem by the method of induction on the order of the group G. Assuming that the theorem is true for all groups of order less than that of G, we shall prove that it is also true for G. We need not worry about starting the induction because the theorem is obviously true for any group of order one.

Now two cases arise :

Case 1. $H_1 = K_1$. In this case after removing G from (1) and (2), we get the remaining series as two composition series for H_1. But the order of H_1 is less than that of G because H_1 is a proper normal subgroup of G. Therefore by our induction hypothesis, the theorem is true for H_1. Since $G/H_1 = G/K_1$, therefore the theorem will remain true if we replace G in each of the series (1) and (2).

Case 2. $H_1 \neq K_1$. By the third law of isomorphism, we have

$$H_1K_1/H_1 \cong K_1/H_1 \cap K_1,$$

and

$$H_1K_1/K_1 \cong H_1/H_1 \cap K_1.$$

Also H_1K_1 is a normal subgroup of G containing H_1. Since H_1 is maximal in G, therefore we must have $H_1K_1=G$.

$$\therefore \quad G/H_1 \cong K_1/D \text{ where } D=H_1 \cap K_1$$

and
$$G/K_1 \cong H_1/D.$$

Now H_1 is maximal in G implies that G/H_1 is simple. Therefore K_1/D is simple and this implies that D is a maximal normal subgroup of K_1. Similarly D is a maximal normal subgroup of H_1.

Let
$$D,\ D_1,\ D_2, \ldots,\ D_t=\{e\}$$
be a composition series for D. Then

$$G,\ H_1,\ D,\ D_1,\ D_2, \ldots,\ D_t=\{e\} \qquad \ldots(3)$$

and
$$G,\ K_1,\ D,\ D_1,\ D_2, \ldots,\ D_t=\{e\} \qquad \ldots(4)$$

are two composition series for G. Let us write the composition quotient groups of (3) and (4) in the order

$$G/H_1,\ H_1/D,\ D/D_1,\ D_1/D_2, \ldots,\ D_{t-1}/D_t \qquad \ldots(5)$$

and
$$K_1/D,\ G/K_1,\ D/D_1,\ D_1/D_2, \ldots,\ D_{t-1}/D_t. \qquad \ldots(6)$$

The quotient groups in (5) and (6) are equal in number and the corresponding quotient groups isomorphic *i.e.*, G/H_1 and K_1/D, H_1/D and G/K_1, D/D_1 and $D/D_1, \ldots$, are isomorphic.

Now (1) and (3) are two composition series for G each having H_1 in the second place. Therefore by case 1, the quotient groups defined by (1) and (3) may be put into $1-1$ correspondence so that the corresponding quotient groups are isomorphic. Similarly the quotient groups defined by (2) and (4) may be put into $1-1$ correspondence so that the corresponding quotient groups are isomorphic. Hence the quotient groups defined by (1) and (2) are equal in number and are isomorphic in some order because the relation of isomorphism in the set of all groups is an equivalence relation. This completes the proof of the theorem.

§ 13. Solvable groups.

Definition. *A group G is said to be solvable if we can find a finite chain of subgroups*

$$G=N_0 \supseteq N_1 \supseteq N_2 \supseteq \cdots \supseteq N_k=(e)$$

such that each N_i is a normal subgroup of N_{i-1} and each quotient group N_{i-1}/N_i is abelian. The above series, then is referred to as *a solvable series for G.* **(Punjab 1968; Mysore 70)**

Normal series of a group. Definition. *A finite sequence of subgroups*

$$G=G_0 \supseteq G_1 \supseteq G_2 \supseteq \cdots \supseteq G_k=(e)$$

of a group G is called a subnormal series of G if G_{i+1} is a normal subgroup of G_i ∀ $i=0, 1, \ldots, k-1$. The quotient groups G_i/G_{i+1} are called the factor groups of the subnormal series. Further if each G_i is a normal subgroup of G itself, then the series is said to be a **normal series** of G.

Solved Examples

Ex. 1. *Show that every abelian group is solvable.*

Solution. Let G be an abelian group. Take $N_0=G$ and $N_1=(e)$. Then $G=N_0 \supseteq N_1=(e)$ is a solvable series for G. Obviously $N_1=(e)$ is a normal subgroup of $N_0=G$ because if a is any element of G, then $a^{-1}ea=a^{-1}a=e \in (e)$.

Further since G is abelian, the quotient group $N_0/N_1=G/(e)$ is also abelian. [Note that every quotient group of an abelian group is abelian]. Hence G is a solvable group.

Ex. 2. *Show that the symmetric group P_3 of degree 3 is solvable.*

Solution. The symmetric group P_3 consists of the six permutations I (identity permutation), (1 2), (2 3), (3 1), (1 2 3) and (1 3 2) on three symbols 1, 2, 3. Let $A_3=\{I, (1\ 2\ 3), (1\ 3\ 2)\}$. Then A_3 is the alternating group of permutations of degree 3. If we take

$$N_0=P_3,\ N_1=A_3, \text{ and } N_2=(I),$$

then $$P_3=N_0 \supseteq N_1 \supseteq N_2=(I)$$

is a solvable series for P_3 as shown below :

We know that A_n is a normal subgroup of P_n. Therefore $A_3=N_1$ is a normal subgroup of $P_3=N_0$. Also (I) is a normal subgroup of N_1. The quotient groups P_3/N_1 and $N_1/(I)$ are of orders 2 and 3 respectively. We know that all groups of orders 2 and 3 are abelian. Therefore the quotient groups P_3/N_1 and $N_1/(I)$ are abelian. Hence $P_3=N_0 \supseteq N_1 \supseteq N_2=(I)$ is a solvable series for P_3 and thus P_3 is solvable.

Ex. 3. *Show that the symmetric group P_4 of degree 4 is solvable.*

Solution. Let A_4 be the alternating group of permutations of degree 4. Then A_4 is a normal subgroup of P_4. Let

$$V_4=\{I, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\}.$$

It can be easily seen that V_4 is a normal subgroup of A_4. If we take

$$N_0=P_4,\ N_1=A_4,\ N_2=V_4 \text{ and } N_3=(I), \text{ then}$$

$$P_4=N_0 \supseteq N_1 \supseteq N_2 \supseteq N_3=(I)$$

is a solvable series for P_4 as shown below :

Obviously (I) is a normal subgroup of N_2. The quotient groups P_4/N_1, N_1/N_2 and N_2/N_3 are of orders 2, 3 and 4 respectively. We know that all groups of orders 2, 3 and 4 are abelian. Hence $P_4=N_0 \supseteq N_1 \supseteq N_2 \supseteq N_3=(I)$ is a solvable seri s for P_4 and thus P_4 is solvable.

Ex. 4. *Prove that a subgroup of a solvable group is solvable.*

Solution. Suppose H is any subgroup of a solvable group G. Let

$$G=G_0 \supseteq G_1 \supseteq G_2 \supseteq \ldots \supseteq G_n=(e)$$

be a solvable series for G. We claim that

$$H=H_0 \supseteq (H \cap G_1) \supseteq (H \cap G_2) \supseteq \ldots \supseteq (H \cap G_n)=(e) \quad \ldots(1)$$

is a solvable series for H. Since for $i=0, 1, \ldots, n-1$, G_{i+1} is normal in G_i, therefore $H_{i+1}=H \cap G_{i+1}$ is normal in $H_i=H \cap G_i$. Let us define a mapping $f : H_i \to G_i/G_{i+1}$ such that

$$f(x)=xG_{i+1} \ \forall \ x \in H_i.$$

Since $H_i \subseteq G_i$, therefore $x \in H_i \Rightarrow x \in G_i$. Consequently the coset xG_{i+1} is an element of the quotient group G_i/G_{i+1} and thus the mapping f is well-defined. Also if $x, y \in H_i$, then

$$f(xy)=(xy)\,G_{i+1} \qquad \text{[by def. of } f\text{]}$$
$$=(xG_{i+1})\,(yG_{i+1}) \qquad [\because \ x, y \in G_i \text{ and } G_{i+1} \text{ is normal in } G_i]$$
$$=f(x)\,f(y).$$

Therefore the mapping f is a homomorphism of H_i into G_i/G_{i+1}. Further

$x \in \ker f \Leftrightarrow f(x)=G_{i+1}$ [Note that the identity of G_i/G_{i+1} is G_{i+1}]

$\Leftrightarrow xG_{i+1}=G_{i+1} \Leftrightarrow x \in G_{i+1}$

$\Leftrightarrow x \in H \cap G_{i+1}$, since $x \in H_i \subseteq H$.

Therefore $\ker f=H \cap G_{i+1}=H_{i+1}$. Hence by the fundamental theorem on homomorphism of groups we have

$$H_i/H_{i+1} \cong f(H_i).$$

But $f(H_i)$ is a subgroup of G_i/G_{i+1} which is abelian. Therefore $f(H_i)$ is also abelian. Consequently H_i/H_{i+1} is also abelian because it is isomorphic to $f(H_i)$. Hence (1) is a solvable series for H and thus H is a solvable group.

Ex. 5. *If G is a group and N is a normal subgroup of G such that both N and G/N are solvable, prove that G is solvable.*

(Mysore 1970)

Solution. It is given that the group G/N is solvable. The identity of this group is N. Let

$$G/N=G_0/N \supseteq G_1/N \supseteq \ldots \supseteq G_{m-1}/N \supseteq G_m/N=(N) \quad \ldots(1)$$

be a solvable series for G/N. Here each G_i is a subgroup of G

containing N. Since G_{i+1}/N is normal in G_i/N, therefore each G_{i+1} is normal in G_i. Also $(G_i/N)/(G_{i+1}/N) \cong G_i/G_{i+1}$. [See theorem 2 page 229]. But each quotient group $(G_i/N)/(G_{i+1}/N)$ is abelian because (1) is a solvable series for G/N. Therefore each quotient group G_i/G_{i+1} is also abelian because it is isomorphic to an abelian group. Further $G_m/N=(N)$ implies $G_m=N$. Also it is given that N is solvable. Let

$$N=N_0 \supseteq N_1 \supseteq N_2 \supseteq \dots \supseteq N_t=(e)$$

be a solvable series for N. Then

$$G=G_0 \supseteq G_1 \supseteq G_2 \supseteq \dots \supseteq G_{m-1} \supseteq N \supseteq N_1 \supseteq N_2 \supseteq \dots \supseteq N_t=(e)$$

is a solvable series for G. Hence G is a solvable group.

Now we shall give an important characteristic property of solvable groups. In this characteristic property we shall use the concept of commutator subgroup of a group which we shall just define.

Commutator subgroup of a group. Definition. *Let G be a group and $a, b \in G$. The element $aba^{-1}b^{-1}$ is called the commutator of the ordered pair (a, b).*

Let $U=\{aba^{-1}b^{-1} : a, b \in G\}$. If G' is the subgroup of G generated by U, then G' is called the commutator subgroup of G.

(Punjab 1970; Lucknow 70 ; Dibrugarh 67 ; Delhi 70)

We recall that if G' is the subgroup of G generated by U, then G' is the smallest subgroup of G containing U. Therefore the commutator subgroup G' of a group G is the smallest subgroup of G containing the set of all commutators in G.

Note. We can also define the commutator of the ordered pair (a, b) as the element $a^{-1}b^{-1}ab$. It will not change the set U of all commutators in G and consequently the commutator subgroup G' will also not change. Note that $a, b \in G \Rightarrow a^{-1}, b^{-1} \in G$. Also $a^{-1}b^{-1}ab$ can be written as $a^{-1}b^{-1}(a^{-1})^{-1}(b^{-1})^{-1}$. Thus we can take $U=\{aba^{-1}b^{-1} : a, b \in G\}$ or we can take

$$U=\{a^{-1}b^{-1}ab : a, b \in G\}.$$

Theorem 1. *Let G' be the commutator subgroup of a group G. Then G is abelian if and only if $G'=\{e\}$, e being the identity element of G*

Proof Let G be any group and let $U=\{aba^{-1}b^{-1} : a, b \in G\}$. If G' is the commutator subgroup of G, then G' is the subgroup of G generated by U *i.e.*, G' is the smallest subgroup of G containing U.

Suppose G is abelian. Then to prove that $G'=\{e\}$. If G is abelian, then $\forall\ a, b \in G$, we have

$$aba^{-1}b^{-1}=abb^{-1}a^{-1}=aea^{-1}=aa^{-1}=e.$$

Thus in this case U consists of only one element *i.e*, e. Now $\{e\}$ is the smallest subgroup of G containing $\{e\}$. Consequently $G'=\{e\}$.

Conversely, suppose that $G'=\{e\}$. Then to prove that G is abelian. Let a, b be any two elements of G. Then $aba^{-1}b^{-1} \in U$. Consequently $aba^{-1}b^{-1} \in G'$. But G' contains only one element *i.e.*, e. Therefore $aba^{-1}b^{-1}=e \Rightarrow (ab)(ba)^{-1}=e \Rightarrow (ab)=[(ba)^{-1}]^{-1}$ $\Rightarrow ab=ba \Rightarrow G$ is abelian.

Theorem 2. *Let G be a group and G′ be the commutator subgroup of G. Then*

(*i*) *G' is normal in G.*

(*ii*) *G/G' is abelian.* **(Meerut 1969 ; Osmania 72; Dibrugarh 67 ; Lucknow 70 ; Delhi 70 ; Punjab 70)**

(*iii*) *If N is any normal subgroup of G, then G/N is abelian if and only if $G' \subseteq N$.* **(Punjab 1968 ; Delhi 70, Lucknow 70)**

(*iv*) *If H is a subgroup of G such that $H \supseteq G'$, then H is a normal subgroup of G.*

Proof. Let $U=\{aba^{-1}b^{-1} : a, b \in G\}$. If G' is the commutator subgroup of G, then G' is the smallest subgroup of G containing U.

(i) Let x be any element of G and c be any element of G'. Then $xcx^{-1}=(xcx^{-1})\,c^{-1}c=(xcx^{-1}c^{-1})\,c \in G$.

$$[\because\ x, c \in G \Rightarrow xcx^{-1}c^{-1} \in G'. \text{ Also } c \in G'.]$$

Hence G' is a normal subgroup of G'.

(ii) Since G' is normal in G, therefore the quotient group G/G' is meaningful. Let a, b be any two elements of G. Then $G'a$, $G'b$ are any two elements of G/G'.

We have $aba^{-1}b^{-1} \in U$

$$\begin{aligned}
&\Rightarrow aba^{-1}b^{-1} \in G' && [\because\ U \subseteq G']\\
&\Rightarrow (ab)(ba)^{-1} \in G' \Rightarrow G'(ab)=G'(ba)\\
&\Rightarrow (G'a)(G'b)=(G'b)(G'a) \Rightarrow G/G' \text{ is abelian.}
\end{aligned}$$

(iii) Let N be any normal subgroup of G. Let $a, b \in G$. Then Na, Nb are any two elements of G/N.

Let G/N be abelian. Then

$$\begin{aligned}
&(Na)(Nb)=(Nb)(Na)\\
&\Rightarrow Nab=Nba \Rightarrow (ab)(ba)^{-1} \in N\\
&\Rightarrow aba^{-1}b^{-1} \in N\\
&\Rightarrow U \subseteq N. && [\because\ aba^{-1}b^{-1} \text{ is any element of } U]
\end{aligned}$$

Thus N is a subgroup of G containing U. Since G' is the smallest subgroup of G containing U, therefore G' must be contained in N *i.e.*, we must have $G' \subseteq N$.

Conversely, let $G' \subseteq N$.

Then $U \subseteq N$ $\quad [\because \; U \subseteq G']$

$$\Rightarrow aba^{-1}b^{-1} \in N$$

$$\Rightarrow (ab)(ba)^{-1} \in N \Rightarrow Nab = Nba$$

$$\Rightarrow (Na)(Nb) = (Nb)(Na) \Rightarrow G/N \text{ is abelian.}$$

(iv) Let g be any element of G and let h be any element of H. Then

$$ghg^{-1} = (ghg^{-1})(h^{-1}h) = (ghg^{-1}h^{-1})h \in H$$

$$[\because \; h \in H. \text{ Also } ghg^{-1}h^{-1} \in G' \text{ and } H \supseteq G' \Rightarrow ghg^{-1}h^{-1} \in H]$$

Hence H is a normal subgroup of G.

Remark. Suppose G' is the commutator subgroup of a group G. Then G' is a group in its own right So we may speak of its commutator subgroup $G^{(2)} = (G')'$. Thus $G^{(2)}$ is a subgroup of G' and hence the subgroup of G generated by all elements $a'\,b'\,(a')^{-1}\,(b')^{-1}$ where $a', b' \in G'$. Also by part (i) of the above theorem 2, $G^{(2)}$ is a normal subgroup of G'. It can also be proved that $G^{(2)}$ is a normal subgroup of G as well. Continuing in this way, we can define higher commutator subgroup $G^{(m)}$ by $G^{(m)} = (G^{(m-1)})'$. This $G^{(m)}$ is called mth commutator subgroup or mth derived subgroup of G. It can be easily shown that each $G^{(m)}$ is a normal subgroup of G. Also by part (ii) of the above theorem 2, each $G^{(m-1)}/G^{(m)}$ is an abelian group.

In terms of higher commutator subgroups of a group G, we shall now give a very important criterion for solvability of G.

A characteristic property of solvable groups.

Theorem 3. *A group G is solvable if and only if $G^{(k)} = (e)$ for some integer k.* **(Meerut 1971)**

Proof. The 'if' part. Let $G^{(k)} = (e)$ for some integer k. Then to prove that G is solvable. Let $N_0 = G$, $N_1 = G'$, $N_2 = G^{(2)}, \ldots,$ $N_k = G^{(k)} = (e)$. Then

$$G = N_0 \supseteq N_1 \supseteq N_2 \supseteq \cdots \supseteq N_k = (e). \qquad \ldots(1)$$

We claim that (1) is a solvable series for G. By part (i) of the preceding theorem 2, $G^{(i)} = (G^{(i-1)})'$ is a normal subgroup of $G^{(i-1)}$ for each i. Therefore N_i is a normal subgroup of N_{i-1} for each i.

Also $\dfrac{N_{i-1}}{N_i} = \dfrac{G^{(i-1)}}{G^{(i)}} = \dfrac{G^{(i-1)}}{(G^{(i-1)})'}$.

But by part (ii) of theorem 2, $\dfrac{G^{(i-1)}}{(G^{(i-1)})'}$ is abelian. Therefore

$\frac{N_{i-1}}{N_i}$ is abelian for each i. Hence (1) is a solvable series for G and thus G is solvable.

The 'only if' part. Let G be a solvable group. Then to prove that $G^{(k)}=(e)$ for some integer k. Let

$$G=N_0 \supseteq N_1 \supseteq N_2 \supseteq \cdots \supseteq N_k=(e) \qquad \ldots(2)$$

be a solvable series for G. Here each N_i is a normal subgroup of N_{i-1} and each N_{i-1}/N_i is an abelian group. Therefore by part (iii) of theorem 2, the commutator subgroup N_{i-1}' of N_{i-1} must be contained in N_i. Thus

$N_1 \supseteq N_0'=G'$,

$N_2 \supseteq N_1' \supseteq (G')'=G^{(2)}$, [Note that $N_1 \supseteq G' \Rightarrow N_1' \supseteq (G')'$]

$N_3 \supseteq N_2' \supseteq (G^{(2)})'=G^{(3)}$,

$N_i \supseteq G^{(i)}$,

$N_k \supseteq G^{(k)}$.

In this way, we see that $(e)=N_k \supseteq G^{(k)}$. Thus $G^{(k)} \subseteq (e)$. But $(e) \subseteq G^{(k)}$ always. Hence $G^{(k)}=(e)$.

Corollary. *Every homomorphic image of a solvable group is solvable.*

Proof. Let G be a solvable group and let G^* be a homomorphic image of G under the homomorphism ψ. Then to prove that G^* is also solvable. It can be easily seen that $(G^*)^{(k)}$ is the image of $G^{(k)}$ under the mapping ψ *i.e.*, $\psi(G^{(k)})=(G^*)^{(k)}$.

Since G is solvable therefore by the preceding theorem 3, we have $G^{(k)}=(e)$ for some integer k. Then

$$(G^*)^{(k)}=\psi\,(G^{(k)})=\psi\,(e)=(e^*),$$

where e^* is the identity of the group G^*. Note that under a homomorphism identity goes to identity. Now by theorem 3,

$$(G^*)^{(k)}=(e^*) \Rightarrow G^* \text{ is solvable.}$$

Ex. *If a group $G\neq(e)$ is solvable, then show that G contains a normal abelian subgroup $H\neq(e)$.*

Solution. In case the group G is abelian, we can take $H=G$ itself and the proof is complete. So let us consider the case when the group G is non-abelian. Since G is a solvable group, therefore by theorem 3, there exists a positive integer k such that $G^{(k)}=(e)$. As G is non-abelian, we cannot have $k=1$. [See theorem 1, page 239]. If k is the least positive integer such that $G^{(k)}=(e)$, let us set $H=G^{(k-1)}$. Then H is a normal subgroup of G and

$H \neq (e)$. Also $H' = (G^{(k-1)})' = G^{(k)} = (e)$. Therefore H is abelian.
[See theorem 1, page 239]

Theorem 4. *Let G denote the symmetric group P_n of degree n where $n \geqslant 5$. Then $G^{(k)}$ contains every 3-cycle of P_n for*

$$k = 1, 2, 3, \ldots.$$

Proof. While proving this theorem, we shall use the following result :

If N is a normal subgroup of a group G, then the commutator subgroup N' of N is also a normal subgroup of G.

We leave the proof of this result as an exercise to the reader. Now we come to the proof of the main theorem.

We claim that if N is a normal subgroup of $G = P_n$ ($n \geqslant 5$) such that N contains every 3-cycle in P_n, then N' must also contain every 3-cycle. For suppose that $f = (1\ 2\ 3)$, $g = (1\ 4\ 5)$ are two three cycles in N. (Note that here $n \geqslant 5$). Then $f g f^{-1} g^{-1}$ as a commutator of elements of N must be a member of N'. But

$$f g f^{-1} g^{-1} = (1\ 2\ 3)\,(1\ 4\ 5)\,(3\ 2\ 1)\,(5\ 4\ 1) = (1\ 5\ 3).$$

Therefore $(1\ 5\ 3) \in N'$. But N' is a normal subgroup of G as mentioned in the beginning of this proof. Therefore if π is any element of P_n, then by the definition of a normal subgroup, $\pi\,(1\ 5\ 3)\pi^{-1}$ must also be in N'. Now let (i_1, i_2, i_3) be any 3-cycle in P_n where i_1, i_2 and i_3 are any three distinct integers in the range from 1 to n. Let us take $\pi \in P_n$ such that $\pi(i_1) = 1$, $\pi(i_2) = 5$, and $\pi(i_3) = 3$.

Then $\pi\,(1\ 5\ 3)\,\pi^{-1} = (i_1, i_2, i_3)$.

[Note that $\pi^{-1}(1) = i_1$, $\pi^{-1}(5) = i_2$, $\pi^{-1}(3) = i_3$. So i_1 goes to 1 under π, then 1 goes to 5 under $(1\ 5\ 3)$ and then 5 goes to i_2 under π^{-1}. Therefore i_1 goes to i_2 under $\pi(1\ 5\ 3)\,\pi^{-1}$. Similarly i_2 goes to i_3 and i_3 goes to i_1 under $\pi\,(1\ 5\ 3)\,\pi^{-1}$. Further if x is any integer other than i_1, i_2 and i_3 in the range from 1 to n, then the image of x under the permutation π cannot be any of 1, 5 and 3 because π is a one-to-one mapping. Therefore x will remain unchanged under the permutation $\pi\,(1\ 5\ 3)\,\pi^{-1}$. Thus we have $\pi\,(1\ 5\ 3)\,\pi^{-1} = (i_1, i_2, i_3)$.]

As mentioned above $\pi\,(1\ 5\ 3)\,\pi^{-1} \in N'$. Therefore $(i_1, i_2, i_3) \in N'$ and thus N' contains every 3-cycle in P_n.

Now $G = P_n$ is definitely normal in G and contains every 3-cycle in P_n. So taking $N = G$, we see that G' contains all 3-cycles. Since G' is normal in G, therefore $(G')' = G^{(2)}$ contains all 3-cycles.

Again since $G^{(3)}$ is normal in G, therefore $G^{(3)}$ contains all 3-cycles. Continuing in this way we obtain that $G^{(k)}$ contains all 3-cycles for arbitrary k.

Theorem 5. *The symmetric group P_n is not solvable for $n \geqslant 5$.*

Proof. Let $G = P_n$ where $n \geqslant 5$. Then by theorem 4, $G^{(k)}$ contains all 3-cycles in P_n for every k. Hence $G^{(k)} \neq (e)$ for any k. Therefore by theorem 3, G is not a solvable group.

§ 14. Direct products.

External direct product. Definition. *Let G_1 and G_2 be any two groups the composition in each being denoted multiplicatively. Then $G_1 \times G_2 = \{(g_1, g_2) : g_1 \in G_1, g_2 \in G_2\}$.*

Let us define a binary operation on $G_1 \times G_2$ denoted multiplicatively as follows :

$(g_1, g_2)\ (h_1, h_2) = (g_1h_1, g_2h_2)$ *where* $g_1, h_1 \in G_1$ *and* $g_2, h_2 \in G_2$.

For this binary operation $G_1 \times G_2$ is a group and this group is called the external direct product of G_1 by G_2. **(Andhra 1975)**

Proof of the fact that $G_1 \times G_2$ is a group for the binary operation we have defined on it.

1. **Closure property.** We have $g_1h_1 \in G_1$ because G_1 is a group. Similarly $g_2h_2 \in G_2$. Thus $(g_1, g_2)\ (h_1, h_2) = (g_1h_1, g_2h_2) \in G_1 \times G_2$.

2. **Associativity.** If $(g_1, g_2), (h_1, h_2), (k_1, k_2) \in G_1 \times G_2$, then $[(g_1, g_2)\ (h_1, h_2)]\ (k_1, k_2) = (g_1h_1, g_2h_2)\ (k_1, k_2) = ([g_1h_1]\ k_1, [g_2h_2]\ k_2)$ $= (g_1\,[h_1k_1], g_2\,[h_2\,k_2]) = (g_1, g_2)\ (h_1k_1, h_2k_2) = (g_1, g_2)\ [(h_1, h_2)(k_1, k_2)]$.

3. **Existence of left identity.** Let e_1, e_2 be the identity elements of G_1, G_2 respectively. If $(g_1, g_2) \in G_1 \times G_2$, then $(e_1, e_2)\ (g_1, g_2) = (e_1g_1, e_2g_2) = (g_1, g_2)$. Therefore (e_1, e_2) is the left identity of $G_1 \times G_2$.

4. **Existence of left inverse.** Let $(g_1, g_2) \in G_1 \times G_2$. Then

$$(g_1^{-1}, g_2^{-1}) \in G_1 \times G_2.$$

Also $(g_1^{-1}, g_2^{-1})\ (g_1, g_2) = (g_1^{-1}\,g_1, g_2^{-1}\,g_2) = (e_1, e_2)$.

$\therefore$ (g_1^{-1}, g_2^{-1}) is the left inverse of (g_1, g_2) in $G_1 \times G_2$.

Hence $G_1 \times G_2$ is a group under the binary operation as defined above.

Theorem 1. *If G_1 and G_2 are groups, then the subsets*

$$G_1 \times \{e_2\} \text{ and } \{e_1\} \times G_2 \text{ of } G_1 \times G_2$$

are normal subgroups of $G_1 \times G_2$ isomorphic to G_1 and G_2 respectively.

Proof. Let (g_1, e_2) and (h_1, e_2) be any two elements of

$$G_1 \times \{e_2\} \text{ where } g_1, h_1 \in G_1.$$

Then $(g_1, e_2)(h_1, e_2)^{-1}$

$$=(g_1, e_2)(h_1^{-1}, e_2^{-1})=(g_1, e_2)(h_1^{-1}, e_2)$$
$$=(g_1h_1^{-1}, e_2e_2)=(g_1h_1^{-1}, e_2).$$

Now $g_1h_1^{-1} \in G_1$ because G_1 is a group.

$$\therefore \quad (g_1h_1^{-1}, e_2) \in G_1 \times \{e_2\}.$$

Hence $G_1 \times \{e_2\}$ is a subgroup of $G_1 \times G_2$.

Now to show that $G_1 \times \{e_2\}$ is normal in $G_1 \times G_2$.

Let (x_1, x_2) be any element of $G_1 \times G_2$ and (g_1, e_2) be any element of $G_1 \times \{e_2\}$.

Then $(x_1, x_2)(g_1, e_2)(x_1, x_2)^{-1}$

$=(x_1, x_2)(g_1, e_2)(x_1^{-1}, x_2^{-1})=(x_1g_1x_1^{-1}, x_2e_2x_2^{-1})$

$=(x_1g_1x_1^{-1}, e_2) \in G_1 \times \{e_2\}$ because $x_1g_1x_1^{-1} \in G_1$.

$\therefore \quad G_1 \times \{e_2\}$ is normal in $G_1 \times G_2$.

Now to show that $G_1 \cong G_1 \times \{e_2\}$.

Let $\phi : G_1 \to G_1 \times \{e_2\}$ defined by

$$\phi(g_1)=(g_1, e_2) \ \forall \ g_1 \in G_1.$$

Obviously ϕ is one-one onto. Also if $g_1, h_1 \in G_1$, then

$\phi(g_1h_1)=(g_1h_1, e_2)=(g_1, e_2)(h_1, e_2)=\phi(g_1)\,\phi(h_1)$.

$$\therefore \quad G_1 \cong G_1 \times \{e_2\}.$$

Similarly we can show that $\{e_1\} \times G_2$ is a normal subgroup of $G_1 \times G_2$ and is isomorphic to G_2.

Theorem 2. *If G_1 and G_2 are groups, then*

(*i*) $G_1 \times \{e_2\} \cap \{e_1\} \times G_2 = \{(e_1, e_2)\}$ *i.e., the identity is the only element common to $G_1 \times \{e_2\}$ and $\{e_1\} \times G_2$.*

(*ii*) *Every element of $G_1 \times \{e_2\}$ commutes with every element of $\{e_1\} \times G_2$.*

(*iii*) *Every element of $G_1 \times G_2$ can be uniquely expressed as the product of an element in $G_1 \times \{e_2\}$ by an element in $\{e_1\} \times G_2$.*

(*iv*) $G_1 \times G_2 \cong G_2 \times G_1$.

Proof. (i) Let $(g, h) \in G_1 \times \{e_2\} \cap \{e_1\} \times G_2$.

Then $(g, h) \in G_1 \times \{e_2\}$ and $(g, h) \in \{e_1\} \times G_2$.

Now $\quad (g, h) \in G_1 \times \{e_2\} \Rightarrow h=e_2$

and $\quad (g, h) \in \{e_1\} \times G_2 \Rightarrow g=e_1$.

$$\therefore \quad (g, h)=(e_1, e_2).$$
$$\therefore \quad G_1 \times \{e_2\} \cap \{e_1\} \times G_2 = \{(e_1, e_2)\}.$$

(ii) Let $(g_1, e_2) \in G_1 \times \{e_2\}$ and $(e_1, g_2) \in \{e_1\} \times G_2$.

Then $(g_1, e_2)(e_1, g_2)=(g_1e_1, e_2g_2)=(g_1, g_2)$

$$=(e_1g_1, g_2e_2)=(e_1, g_2)(g_1, e_2).$$

(iii) Let $(g_1, g_2) \in G_1 \times G_2$.

Then $(g_1, g_2)=(g_1e_1, e_2g_2)=(g_1, e_2)(e_1, g_2)$.

Thus (g_1, g_2) can be expressed as a product of a member of $G_1 \times \{e_2\}$ by a member of $\{e_1\} \times G_2$.

If possible, let there be another representation

$$(g_1, g_2)=(h_1, e_2)\,(e_1, h_2).$$

Then $(g_1, g_2)=(h_1e_1, e_2h_2)=(h_1, h_2)$.

$\therefore$ $g_1=h_1$, $g_2=h_2$ by the equality of ordered pairs.

$\therefore$ The representation is unique.

(iv) Define a mapping $f : G_1 \times G_2 \to G_2 \times G_1$ by

$$f(g_1, g_2)=(g_2, g_1).$$

Obviously f is one-one and onto.

Also if $(g_1, g_2), (h_1, h_2) \in G_1 \times G_2$, then

$$f[(g_1, g_2)\,(h_1, h_2)]=f(g_1h_1, g_2h_2)=(g_2h_2, g_1h_1)$$
$$=(g_2, g_1)\,(h_2, h_1)=f(g_1, g_2)\,f(h_1, h_2).$$

$\therefore$ $G_1 \times G_2 \cong G_2 \times G_1$.

Internal direct products.

Definition. *Suppose H and K are subgroups of a group G. Then we say that G is an internal direct product of H and K if*

(*i*) *every element of H commutes with every element of K.*

(*ii*) *every element of G is uniquely expressible as a product of an element of H by an element of K.* **(Andhra 1975)**

Theorem 1. *Suppose a group G is an internal direct product of its subgroups H and K. Then*

(*i*) *H and K have only the identity in common.*

(*ii*) *G is isomorphic to the external direct product of H by K.*

Proof. (i) Let $x \in H$, $x \in K$.

Since H and K are subgroups, therefore

$$x^{-1} \in H,\ x^{-1} \in K.$$

Since G is an internal direct product of H and K, therefore every element g in G can be uniquely expressed in the form

$$g=hk \text{ where } h \in H,\ k \in K.$$

Also we can write $g=(hx)\,(x^{-1}k)$ where $hx \in H$, $x^{-1}k \in K$.

Since the expression for g is unique, therefore

$$hx=h \Rightarrow hx=he \Rightarrow x=e.$$

$\therefore$ e is the only element common to both H and K.

(ii) Let g be an arbitrary element of G.

Then $g=hk$ where h is a unique element of H and k is a unique element of K.

Consider the function $\phi : G \to H \times K$ defined by

$$\phi(g)=(h, k)\ \forall\ g \in G.$$

ϕ is one-one. Let $g_1=h_1k_1$, $g_2=h_2k_2$ be any two elements of G where $h_1, h_2 \in H$ and $k_1, k_2 \in K$.

We have $\phi(g_1)=\phi(g_2) \Rightarrow (h_1, k_1)=(h_2, k_2)$

$\Rightarrow h_1=h_2, k_1=k_2 \Rightarrow h_1k_1=h_2k_2 \Rightarrow g_1=g_2.$

$\therefore$ ϕ is $1-1$.

ϕ is onto $H \times K$. Let (h, k) be any element of $H \times K$. Then $hk \in G$ and we have $\phi(hk)=(h, k)$.

Therefore ϕ is onto $H \times K$.

ϕ preserves compositions in G and $H \times K$.

Let $g_1=h_1k_1$, $g_2=h_2k_2$ be any two elements of G. Then

$$\begin{aligned}\phi(g_1g_2) &= \phi(h_1k_1h_2k_2)\\ &= \phi(h_1h_2k_1k_2) && [\because \text{ every element of } H \text{ commutes with every element of } K]\\ &= (h_1h_2, k_1k_2) && [\because h_1h_2 \in H, k_1k_2 \in K]\\ &= (h_1, k_1)(h_2, k_2) = \phi(h_1k_1)\,\phi(h_2k_2) = \phi(g_1)\,\phi(g_2).\end{aligned}$$

$\therefore$ $G \cong H \times K$.

Note. If G is an internal direct product of its subgroups H and K, then G is isomorphic to the external direct product of H by K. On account of this isomorphism we shall identify the internal direct product with the external direct product. If G is the internal direct product of its subgroups H and K we shall write $G=H \times K$.

Theorem 2. *If H, K are two subgroups of a group G such that $G=H \times K$, then H, K are normal subgroups of G, and*

$$G/H \cong K \text{ and } G/K \cong H.$$

Proof. Let $g \in G$. Then $g=hk$ where h is a unique element of H and k is a unique element of K. So consider the mapping $\phi : G \to H$ defined by $\phi(g)=\phi(hk)=h$. Obviously ϕ is onto H.

Let $g_1=h_1k_1$, $g_2=h_2k_2$ be any two elements of G where h_1, h_2 are unique elements of H and k_1, k_2 are unique elements of K. Then

$$\begin{aligned}\phi(g_1g_2) &= \phi(h_1k_1h_2k_2)\\ &= \phi(h_1h_2k_1k_2) && [\because \text{ every element of } H \text{ commutes with every element of } K]\\ &= h_1h_2 && [\because h_1h_2 \in H, k_1k_2 \in K]\\ &= \phi(h_1k_1)\,\phi(h_2k_2) = \phi(g_1)\,\phi(g_2).\end{aligned}$$

$\therefore$ ϕ is a homomorphism of G onto H. The kernel of ϕ consists of all those elements of G which are mapped by ϕ on the identity of H *i.e.*, which are of the form $ek=k$, $k \in K$.

Thus K is the kernel of ϕ.

$\therefore$ K is a normal subgroup of G.

Also by the fundamental theorem on homomorphism of groups, we have $G/K \cong H$.

Similarly we can show that H is a normal subgroup of G and

$$G/H \cong K.$$

Theorem 3. *A group G is the direct product of its two subgroups H and K if and only if*

(i) H and K are normal subgroups of G,

(ii) $H \cap K=\{e\}$, and (iii) $G=HK$. **(Andhra 1975)**

Proof. Suppose the conditions (i), (ii) and (iii) hold. Then to show that $G=H\times K$.

Let $h \in H, k \in K$. Consider the element $h^{-1}k^{-1}hk$.

Since H is normal in G, therefore $k^{-1}hk \in H$.

Also $h^{-1} \in H$. Therefore $h^{-1}k^{-1}hk \in H$.

Again K is normal in G. Therefore $h^{-1}k^{-1}h \in K$.

Also $k \in K$. Therefore $h^{-1}k^{-1}hk \in K$.

$\therefore \quad h^{-1}k^{-1}hk \in H \cap K$.

But $\quad H \cap K=\{e\}$.

$\therefore \quad h^{-1}k^{-1}hk=e \Rightarrow hk=kh$.

Hence every element of H commutes with every element of K.

Let $x \in G$. Then by the condition (iii) there exists $h \in H$ and $k \in K$ such that $x=hk$.

If possible let $x=h_1k_1$ where $h_1 \in H$, $k_1 \in K$.

Then $hk=h_1k_1 \Rightarrow h_1^{-1}h=k_1k^{-1}$.

But $h_1^{-1}h \in H$ and $k_1k^{-1} \in K$.

$\therefore \quad h_1^{-1}h=k_1k^{-1} \in H \cap K$.

But $\quad H \cap K=\{e\}$.

$\therefore \quad h_1^{-1}h=k_1k^{-1}=e$.

$\therefore \quad h_1=h, \; k_1=k$.

Thus the expression $x=hk$ for x is unique. Hence $G=H\times K$.

Conversely let $G=H\times K$. Then to show that the conditions (i), (ii) and (iii) are satisfied.

Let a be any element of H and x be any element of G. Then there exist $h \in H$, $k \in K$ such that $x=hk$.

We have $xax^{-1}=(hk)\, a\, (hk)^{-1}=hk\, ak^{-1}h^{-1}$

$=hak\, k^{-1}h^{-1}$ $\quad$ [$\because$ $k \in K$ and $a \in H \Rightarrow k$ commutes with a]

$=hah^{-1} \in H$.

$\therefore$ H is a normal subgroup of G. Similarly K is a normal subgroup of G.

Now we establish condition (ii).

Suppose if possible $b \in H \cap K$ and $b \neq e$. Then $b=be=eb$.

This shows that there exist at least two different ways of expressing an element $b \in G$ as product of an element of H with an element of K. But this contradicts the assumption that

$$G = H \times K.$$

Therefore $b = e$. Hence $H \cap K = \{e\}$.

Finally $HK \subseteq G$. Also $x \in G \Rightarrow x = hk$ where $h \in H, k \in K$. Thus $x \in G \Rightarrow x \in HK$. Therefore $G \subseteq HK$. Hence $G = HK$.

Thus the theorem is completely established.

§ 15. Some important theorems on finite groups.

Theorem 1. Cauchy's theorem for abelian groups. *Suppose G is a finite abelian group and $p \mid o(G)$ i.e. p is a divisor of $o(G)$ where p is a prime number. Then there is an element $a \neq e \in G$ such that $a^p = e$.* **(I.A.S. 1971; Kanpur 71)**

Proof. We shall prove the theorem by induction on the order of G. Assuming that the theorem is true for abelian groups of order less than that of G, we shall prove that it is also true for G. To start the induction we note that the theorem is vacuously true for groups of order one.

If G has no proper subgroups, then G must be of prime order because every group of composite order possesses proper subgroups. But p is prime and $p \mid o(G)$, therefore $o(G)$ must be equal to p. Also every group of prime order is cyclic. Therefore each element $a \neq e$ of G will be a generator of G. Thus G has $p-1$ elements $a \neq e$ such that $a^p = a^{o(G)} = e$.

So now suppose that G has a proper subgroup H *i.e.*,

$$H \neq \{e\} \text{ and } H \neq G. \quad \text{If } p \mid o(H),$$

then by our induction hypothesis the theorem is true for H because H is an abelian group and $o(H) < o(G)$. Therefore $\exists$ an element $b \in H$, $b \neq e$ such that $b^p = e$. But $b \in H \Rightarrow b \in G$ because $H \subseteq G$. Thus $\exists$ an element $b \in G$, $b \neq e$ such that $b^p = e$.

So let us suppose that p is not a divisor of $o(H)$. Since G is abelian, therefore H is a normal subgroup of G and so G/H is a quotient group. Since G is abelian, therefore G/H is also abelian. Also we have $o(G/H) = \dfrac{o(G)}{o(H)} < o(G)$ because $o(H) > 1$.

Since $p \mid o(G)$ and p is not a divisor of $o(H)$, therefore p is a divisor of $o(G)/o(H)$. Hence by our induction hypothesis the theorem is true for the group G/H.

Remembering that H is the identity element of G/H we deduce that $\exists$ an element c in G and $Hc \neq H$ in G/H such that $(Hc)^p = H$.

Now Hc is not equal to the identity element H of the quotient group G/H and p is a prime. Therefore $(Hc)^p=H$ implies that in the quotient group G/H we have $o\ (Hc)=p$.

Also $(Hc)^p=H \Rightarrow Hc^p=H \Rightarrow c^p \in H$. Therefore by a corollary to Lagrange's theorem, we have

$$(c^p)^{o(H)}=e \text{ (i.e., the identity of } H).$$

$\therefore \quad (c^{o(H)})^p=e$ or $d^p=e$ if we put $d=c^{o(H)}$.

This d will be the required element of G if we show that $d \neq e$. If possible, let $d=e$. Then $(Hc)^{o(H)}=Hc^{o(H)}=Hd=He=H$. But in the quotient group G/H, we have $o\ (Hc)=p$. Therefore $(Hc)^{o(H)}=H$ (*i.e.*, the identity of G/H) implies that p must be a divisor of $o(H)$ which is a contradiction. So we cannot have $d=e$. Thus $d \neq e$ and $d^p=e$. Therefore we have completed the induction and this proves the required result.

Theorem 2. Cauchy's theorem. *Suppose G is a finite group and $p \mid o\ (G)$ where p is a prime number. Then there is an element a in G such that $o\ (a)=p$.* (Meerut 1980)

Proof. We shall prove the theorem by induction on $o\ (G)$. Assuming that the theorem is true for groups of order less than that of G, we shall prove that it is also true for G. To start the induction we note that the theorem is vacuously true for groups of order one.

If there exists a subgroup $H \neq G$ of G such that $p \mid o\ (H)$, then by our induction hypothesis the theorem is true for H because $o\ (H) < o\ (G)$. Therefore there exists an element $a \in H$ such that $o\ (a)=p$. But $a \in H \Rightarrow a \in G$ because $H \subset G$. Therefore there exists an element $a \in G$ such that $o\ (a)=p$.

So let us now assume that p is not a divisor of the order of any proper subgroup of G. Let Z be the centre of G. We write the class equation for G in the form :

$$o\ (G)=o\ (Z)+\sum_{a \notin Z} \frac{o\ (G)}{o\ [N\ (a)]} \qquad \ldots(1)$$

[See theorem 6, page 201]

Now $N\ (a)$ is a subgroup of G. If $a \notin Z$, then $N\ (a) \neq G$ and so p is not a divisor of $o\ [N\ (a)]$. But $p \mid o\ (G)$. Therefore

$$p \mid \frac{o\ (G)}{o\ [N(a)]} \text{ if } a \notin Z.$$

$$\therefore \quad p \mid \sum_{a \notin Z} \frac{o\ (G)}{o\ [N\ (a)]}.$$

But $p \mid o(G)$. Therefore $p \mid \left[o(G) - \sum\limits_{a \notin Z} \frac{o(G)}{o[N(a)]} \right]$. Then from (1), we conclude that $p \mid o(Z)$. Thus Z is a subgroup of G and the order of Z is divisible by p. But according to our assumption p is not a divisor of the order of any proper subgroup of G. Consequently $Z=G$. But then G is abelian. Therefore by Cauchy's theorem for abelian groups there exists an element in G of order p.

This completes the proof of the theorem.

Example 1. *Show that if p is a prime number, then any group G of order 2p has a normal subgroup of order p.* **(Meerut 1969)**

Solution. We have $o(G)=2p$.

Since p is prime and $p \mid o(G)$, therefore by Cauchy's theorem G has an element a of order p. Then the cyclic group

$H=\{a, a^2, \ldots, a^p=e\}$ is a subgroup of G of order p.

The index of H in $G = \frac{o(G)}{o(H)} = \frac{2p}{p} = 2$.

But we know that every subgroup of G having index 2 in G is normal in G. Therefore H is normal in G.

Example 2. *Prove that every abelian group of order 6 is cyclic.*

Solution. Let G be an abelian group of order 6. Since the prime integers 3 and 2 are both divisors of $o(G)$, therefore by Cauchy's theorem for abelian groups there exist elements a and b in G such that $o(a)=3$ and $o(b)=2$. We shall show that $o(ab)=6$ and consequently G will be a cyclic group generated by ab.

We have $b^{-1} \neq a$, since $o(b^{-1}) = o(b) = 2$ while $o(a)=3$. Thus $ab \neq e$. Now $(ab)^2 = a^2b^2 = a^2e = a^2 \neq e$, since $o(a)=3$. Again $(ab)^3 = a^3b^3 = eb^3 = b^3 = b^2b = eb = b \neq e$. Therefore we must have $o(ab) > 3$. But $o(ab)$ must be a divisor of $o(G)$ *i.e.*, 6. So $o(ab)$ can neither be 4 nor it can be 5. Hence we must have $o(ab)=6$ and consequently G is cyclic.

Definition. *Suppose G is a finite group and* $o(G)=p^m n$ *where p is a prime number and p is not a divisor of n. Then a subgroup H of G is said to be a p-Sylow subgroup of G iff* $o(H)=p^m$.

Theorem 3. Sylow's theorem. *Suppose G is a group of finite order and p is a prime number. If* $p^m \mid o(G)$ *and* p^{m+1} *is not a divisor of* $o(G)$, *then G has a subgroup of order* p^m.

(Kanpur 1980 ; I.A.S. 72 ; Vikram 76 ; Calicut 75)

Proof. We shall prove the theorem by induction on $o(G)$.

Assuming that the theorem is true for groups of order less than that of G, we shall show that it is also true for G. To start the induction we see that the theorem is obviously true if $o\ (G)=1$.

Let $o\ (G)=p^m n$ where p is not a divisor of n. If $m=0$, the theorem is obviously true. If $m=1$, the theorem is true by Cauchy's theorem. So let $m>1$. Then G is a group of composite order and so G must possess a subgroup H such that $H \neq G$.

If p is not a divisor of $\dfrac{o\ (G)}{o\ (H)}$, then $p^m \mid o\ (H)$ because

$$o\ (G)=p^m n=o\ (H).\ \frac{o\ (G)}{o\ (H)}.$$

Also p^{m+1} cannot be a divisor of $o\ (H)$ because then p^{m+1} will be a divisor of $o\ (G)$ of which $o\ (H)$ is a divisor. Further $o\ (H) < o\ (G)$. Therefore by our induction hypothesis, the theorem is true for H. Therefore H has a subgroup of order p^m and this will also be a subgroup of G. So let us assume that for every subgroup H of G where $H \neq G$, p is a divisor of $\dfrac{o\ (G)}{o\ (H)}$. Consider the class equation,

$$o\ (G)=o\ (Z)+\sum_{a \notin Z} \frac{o\ (G)}{o\ [N\ (a)]} \qquad \ldots(1)$$

Since $a \notin Z \Rightarrow N\ (a) \neq G$, therefore according to our assumption p is a divisor of $\displaystyle\sum_{a \notin Z} \frac{o\ (G)}{o\ [N\ (a)]}$. Also $p \mid o\ (G)$.

Therefore from (1), we conclude that p is a divisor of $o\ (Z)$. Then by Cauchy's theorem, Z has an element b of order p. Z is the centre of G. Also $N=\{b\}$ is a cyclic subgroup of Z of order p. Therefore N is a cyclic subgroup of G of order p. Since $b \in Z$, therefore N is a normal subgroup of G of order p.

[Ex. 7 on page 208 after § 4]

Now consider the quotient group $G'=G/N$.

We have $\quad o\ (G')=o\ (G)/o\ (N)=p^m n/p=p^{m-1}\ n$.

Thus $o\ (G') < o\ (G)$. Also $p^{m-1} \mid o\ (G')$ but pG is not a divisor of $o\ (G')$. Therefore by our induction hypothesis M' has a subgroup, say S' of order p^{m-1}. We know that the natural mapping $\phi : G \rightarrow G/N$ defined by $\phi(x)=Nx \ \forall \ x \in G$ is a homomorphism of G onto G/N with kernel N. Let $S=\{x \in G : \phi(x) \in S'\}$.

Then S is a subgroup of G and $S' \cong S/N$. [See theorem 4 of § 10]

$$\therefore \quad o(S')=o(S/N)=\frac{o(S)}{o(N)}.$$

Therefore $o(S)=o(S').o(N)=p^{m-1}p=p^m$.

Thus S is a subgroup of G of order p^m.

This completes the proof of the theorem.

Solved Examples

Ex. 1. *If H is a p-Sylow subgroup of G and $x \in G$, then $x^{-1}Hx$ is also a p-Sylow subgroup of G.*

Solution. Suppose G is a finite group and $o(G)=p^m n$ where p is a prime number and p is not a divisor of n. If H is a p-Sylow subgroup of G, then $o(H)=p^m$.

Let $x \in G$ be arbitrary. Then $x^{-1}Hx$ will be a p-Sylow subgroup of G if $x^{-1}Hx$ is a subgroup of G and if $o(x^{-1}Hx)=p^m$.

First we shall prove that $x^{-1}Hx$ is a subgroup of G. Let $x^{-1}h_1x$, $x^{-1}h_2x$ be any two elements of $x^{-1}Hx$. Then $h_1, h_2 \in H$. Also we have

$$(x^{-1}h_1x)(x^{-1}h_2x)^{-1}=x^{-1}h_1xx^{-1}h_2^{-1}(x^{-1})^{-1}=x^{-1}h_1eh_2^{-1}x$$
$$=x^{-1}h_1h_2^{-1}x$$

$\in x^{-1}Hx$ since $h_1h_2^{-1} \in H$, H being a subgroup of G.

$\therefore$ $x^{-1}Hx$ is a subgroup of G.

Now let ψ be a mapping from H into $x^{-1}Hx$ defined as

$$\psi(h)=x^{-1}hx \ \forall \ h \in H.$$

ψ **is onto.** Let $x^{-1}hx$ be any element of $x^{-1}Hx$. Then $h \in H$ and we have $\psi(h)=x^{-1}hx$. Therefore ψ is onto.

ψ **is one-one.** Let $h_1, h_2 \in H$. Then

$\psi(h_1)=\psi(h_2) \Rightarrow x^{-1}h_1x=x^{-1}h_2x$

$\Rightarrow h_1=h_2$ [by cancellation laws]

$\Rightarrow \psi$ is one-one.

Thus ψ is a one-to-one correspondence between the elements of H and the elements of $x^{-1}Hx$. Therefore $o(x^{-1}Hx)=o(H)=p^m$. Hence $x^{-1}Hx$ is a p-Sylow subgroup of G.

Ex. 2. *If a group G has only one p-Sylow subgroup H, then H is normal in G.*

Solution. Suppose a group G has only one p-Sylow subgroup H. Let x be any element of G. Then by previous exercise, $x^{-1}Hx$ is also a p-Sylow subgroup of G. But H is the only p-Sylow subgroup of G. Therefore

$x^{-1}Hx=H \ \forall \ x \in G \Rightarrow H$ is a normal subgroup of G.

4

Rings

§ 1. So far we have studied **group** which is an algebraic structure equipped with one binary operation. In this chapter we shall study **ring** which is an algebraic structure equipped with two binary operations.

Ring. Definition. **(I A.S. 1971; Garhwal 76; Guru Nanak 75; Marathwada 70; Vikram 78; Sambalpur 77; Kumayun 77; Gujrat 70; Meerut 81, 84)**

Suppose R is a non-empty set equipped with two binary operations called addition and multiplication and denoted by '+' *and* '.' *respectively i e., for all* $a, b \in R$ *we have* $a+b \in R$ *and* $a\, b \in R$. *Then this algebraic structure* $(R, +, .)$ *is called a* ring, *if the following postulates are satisfied :*

1. *Addition is associative, i.e.,*
$(a+b)+c=a+(b+c) \ \forall\ a, b, c \in R.$
2. *Addition is commutative, i.e.,* $a+b=b+a \ \forall\ a, b \in R.$
3. *There exists an element denoted by* 0 *in R such that*
$$0+a=a \ \forall\ a \in R.$$
4. *To each element a in R there exists an element −a in R such that* $(-a)+a=0$
5. *Multiplication is associative, i.e.,*
$$a\,(b.c)=(a.b)\,c \ \forall\ a, b, c \in R.$$
6. *Multiplication is distributive with respect to addition, i.e , for all a, b, c, in R,*

$$\left.\begin{aligned} a\,(b+c)&=a.b+a.c \\ \text{and} \quad (b+c)\,a&=b.a+c.a \end{aligned}\right\} \begin{aligned} &\textit{Left distributive law} \\ &\textit{Right distributive law.} \end{aligned}$$

Since addition is commutative in R, therefore we shall have $0 \in R$ such that $0+a=a=a+0 \ \forall\ a \in R$.

Also if $a \in R$, then we shall have $(-a)+a=0=a+(-a)$.

Thus R will be an abelian group with respect to addition. The element $0 \in R$ will be the additive identity. It is called the **zero element** of the ring. Since in a group the identity element is unique, therefore every ring will possess a unique zero element and it will be the identity element for addition composition. We shall always denote this element by the symbol 0.

Students should not confuse it with the number 0. It is a symbol which will represent the additive identity of the ring.

In a ring every element will possess a unique inverse for addition composition. We shall denote the additive inverse of a by the symbol $-a$.

We shall define $a-b=a+(-b)$..

The equation $a+x=b$ will have a unique solution in R and it will be $x=b-a$. Obviously $a+(b-a)=a+[b+(-a)]$
$=a+[(-a)+b]=[a+(-a)]+b=[(-a)+a]+b=0+b=b$.

Similarly the equation $y+a=b$ will have a unique solution in R and it will be $y=b-a$.

Both the cancellation laws will hold good for addition in R *i.e.*, for all a, b, c in R

$$a+b=a+c \Rightarrow b=c \text{ and } b+a=c+a \Rightarrow b=c.$$

If in a ring we have $a+b=0$, then $a=-b$ and $b=-a$.

Ring with unity. *If in a ring R there exists an element denoted by* 1 *such that* $1.a=a=a.1 \ \forall \ a \in R$, *then R is called a ring with unit element. The element* $1 \in R$ *is called the unit element of the ring.* Obviously 1 is the multiplicative identity of R. Thus if a ring possesses multiplicative identity, then it is a ring with unity.

Commutative Ring. *If in a ring R, the multiplication composition is also commutative i.e., if we have* $a.b=b.a \ \forall \ a, b \in R$, *then R is called a commutative ring.* **(Meerut 1976, Nagarjuna 78)**

Note. In future we shall denote the multiplication composition in a ring R not by the symbol '.' but by multiplicative notation. Thus we shall write ab in place of $a.b$.

§ 2. Elementary Properties of a Ring.

Theorem. *If R is a ring, then for all* $a, b, c \in R$

(i) $a0=0a=0$. **(Meerut 1970; Sagar 77)**

(ii) $a(-b)=-(ab)=(-a)b$. **(Sambalpur 1977; Bombay 70)**

(iii) $(-a)(-b)=ab$. **(Banaras 1968; Sagar 77)**

(iv) $a(b-c)=ab-ac$. **(Meerut 1970; Vikram 78)**

(v) $(b-c)a=ba-ca$. **(Banaras 1968)**

Proof. (i) We have

$$a0=a(0+0) \qquad [\because 0+0=0]$$
$$=a0+a0. \qquad [\text{by left distributive law}]$$

$\therefore \quad 0+a0=a0+a0 \qquad [\because \; a0 \in R \text{ and } 0+a0=a0]$

Now R is a group with respect to addition, therefore applying right cancellation law for addition in R we get $0=a0$.

Similarly we have $0a=(0+0)\,a$

$=0a+0a.$ [by right distributive law]

$\therefore \quad 0+0a=0a+0a \qquad [\because \; 0+0a=0a].$

Applying right cancellation law for addition in R, we get $0=0a$.

(ii) We have $a\,[(-b)+b]=a0 \qquad [\because \; -b+b=0]$

$\Rightarrow a\,(-b)+ab=0$ [by using left distributive law and the result (i)]

$\Rightarrow a\,(-b)=-(ab)$, since in a ring $a+b=0 \Rightarrow a=-b$.

Similarly we have $(-a+a)\,b=0b$

$\Rightarrow (-a)\,b+ab=0 \Rightarrow (-a)\,b=-(ab)$, since in a ring $a+b=0 \Rightarrow a=-b$.

(iii) We have $(-a)\,(-b)=-[(-a)\,b]$, since $a\,(-b)=-(ab)$

$=-[-(ab)]$, since $(-a)\,b=-(ab)$

$=ab$, since R is a group with respect to addition and in a group we have $-(-a)=a$.

(iv) We have $a\,(b-c)=a\,[b+(-c)]$

$=ab+a\,(-c)$ [left distributive law]

$=ab+[-(ac)] \qquad [\because \; a\,(-c)=-(ac)]$

$=ab-ac.$

(v) We have

$(b-c)\,a=[b+(-c)]\,a$

$=ba+(-c)\,a$ [right distributive law]

$=ba+[-(ca)]=ba-ca.$

§ 3. Integral Multiples of the elements of a ring.

Suppose R is a ring and $a \in R$ If m is a positive integer, then we define $ma=a+a+\ldots$upto m terms.

Also we define $0a=0$. Here 0 on the left hand side is the integer zero and 0 on the right hand side is the zero element (additive identity) of the ring.

If m is a positive integer, then $-m$ is a negative integer. We define $(-m)\,a=-(ma)=-(a+a+a\ldots\text{upto } m \text{ terms})$

$=(-a)+(-a)+(-a)+\ldots\text{upto } m \text{ terms}=m\,(-a).$

If m and n are any integers, we can prove that

$ma+na=(m+n)\,a$, $m\,(na)=(mn)\,a$, and $(ma)\,(na)=(mn)\,a^2$.

Also we can prove that for any integer m, we have

$m\,(a+b)=ma+mb \;\forall\; a, b \in R.$

The students should not confuse that this is the distributive law we assumed in the postulates for a ring. It is not so. Here m is an integer and a, b are elements of R.

§ 4. Examples of Rings.

Example 1. *The set R consisting of a single element* 0 *with two binary operations defined by* $0+0=0$ *and* $0.0=0$ *is a ring. This ring is called the* **null ring** *or the* **zero ring.**

Example 2. *The set* I *of all integers is a ring with respect to addition and multiplication of integers as the two ring compositions. This ring is called the* **ring of integers.** **(Garhwal 1976)**

Solution. As in groups, we should first prove that I is an abelian group with respect to addition of integers.

Further we observe that

(i) The product of two integers is also an integer. Therefore I is closed with respect to addition of integers.

(ii) Multiplication of integers is an associative composition.

(iii) Multiplication of integers is distributive with respect to addition of integers *i.e.*, if a, b, c are any elements of I, then

$$a(b+c)=ab+ac \text{ and } (b+c)\,a=ba+ca.$$

Therefore I is a ring with respect to addition and multiplication of integers. The integer 0 is the zero element of this ring. Also the multiplicative identity exists and is the integer 1. We have $1a=a=a1 \ \forall \ a \in \mathbf{I}$. Thus the ring of integers is a ring with unity. The integer 1 is the unit element of this ring.

The multiplication of integers is a commutative composition. Therefore it is also a commutative ring.

Example 3. *The set* 2I *of all even integers is a commutative ring without unity, the addition and multiplication of integers being the two ring compositions.*

Example 4 *The set* Q *of all rational numbers is a commutative ring with unity, the addition and multiplication of rational numbers being the two ring compositions.*

(Meerut 1976; Kanpur 80; Kumayon 77)

Example 5. *The set* **R** *of all real numbers is a commutative ring with unity, the addition and multiplication of real numbers being the two ring compositions.*

Example 6. *The set* C *of all complex numbers is a commutative ring with unity, the addition and multiplication of complex numbers being the two ring compositions.*

Note. The students should themselves write the complete solutions of examples 3, 4, 5 and 6 with the help of example 2 of the ring of integers and the corresponding questions of groups.

Example 7. *The set M of all $n \times n$ matrices with their elements as real numbers (rational numbers, complex numbers, integers) is a non-commutative ring with unity, with respect to addition and multiplication of matrices as the two ring compositions.*

Solution. We know that the sum and product of two $n \times n$ matrices with their elements as real numbers are again $n \times n$ matrices with their elements as real numbers. Therefore M is closed with respect to addition and multiplication of matrices.

Further we observe that

(i) $A+(B+C)=(A+B)+C \ \forall \ A, B, C \in M$, since the addition of matrices is an associative composition.

(ii) $A+B=B+A \ \forall \ A, B \in M$, since the addition of matrices is commutative.

(iii) If O is the null matrix of the type $n \times n$, then $O \in M$ and we have $O+A=A \ \forall \ A \in M$.

(iv) To each matrix $A \in M$ there exists a matrix $-A \in M$ such that $(-A)+A=O$ (null matrix).

(v) $(AB)\,C=A\,(BC)$, $\forall \ A, B, C \in M$, since multiplication of matrices is associative.

(vi) $A\,(B+C)=AB+AC$,

and $(B+C)\ A=BA+CA \ \forall \ A, B, C \in M$, since matrix multiplication is distributive with respect to matrix addition.

Hence M is a ring with respect to the two given compositions. The null matrix O of the type $n \times n$ is the zero element of this ring *i.e.*, $O=0$.

Since the multiplication of matrices is not in general commutative, therefore the ring is a non-commutative ring. $[n>1]$

Finally if I be the unit matrix of the type $n \times n$, then $I \in M$ and we have $IA=A=AI \ \forall \ A \in M$. Therefore the matrix I is the multiplicative identity. Thus the ring is with unity and the matrix I is the unity element of the ring *i.e.*, $I=1$.

Example 8. *The set $R=\{0, 1, 2, 3, 4, 5\}$ is a commutative ring with respect to '$+_6$' and '$\times_6$' as the two ring compositions.*

Solution. As we have proved in groups, we should first prove that R is an abelian group with respect to '$+_6$'.

Now we form the composition table for R for the composition $\times_6$.

$\times_6$	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1.

From the composition table we see that R is closed with respect to the composition '$\times_6$'.

Also we know that '$\times_6$' is an associative composition in R *i.e.*, $a \times_6 (b \times_6 c) = (a \times_6 b) \times_6 c \ \forall \ a, b, c \in R$.

Further '$\times_6$' is distributive in R with respect to '$+_6$'. If a, b, c are any elements of R, then

$a \times_6 (b +_6 c) = a \times_6 (b+c)$ $\qquad [\because \ b+c \equiv b +_6 c \pmod 6]$

= least non-negative remainder when $a(b+c)$ is divided by 6

= least non-negative remainder when $ab+ac$ is divided by 6

$= (ab) +_6 (ac)$

$= (a \times_6 b) +_6 (ac)$ $\qquad [\because \ a \times_6 b \equiv ab \pmod 6]$

$= (a \times_6 b) +_6 (a \times_6 c)$ $\qquad [\because \ a \times_6 c \equiv ac \pmod 6]$

Similarly we can prove that $(b +_6 c) \times_6 a = (b \times_6 a) +_6 (c \times_6 a)$.

$\therefore$ R is a ring with respect to the given compositions. Since '$\times_6$' is a commutative composition in R as is clear from the composition table also, therefore R is a commutative ring. Also 1 is the identity element for the composition '$\times_6$'. Therefore R is a ring with unity. The integer 0 is the zero element of this ring.

Important. We see that in this ring R neither 2 nor 3 is equal to the zero element of the ring. But $2 \times_6 3 = 0$ (zero element of the ring). Thus in a ring it is possible that the product of two non-zero elements is equal to the zero element. Also the number of elements in R is finite. Therefore this is an example of a finite ring.

§ 5. Some special types of rings. Rings with or without zero divisors.

We have proved that in any ring R, if 0 is the additive identity *i.e.*, the zero element of the ring, then $0a=a0=0\ \forall\ a\in R$. However there are rings in which it is possible that $ab=0$ when neither $a=0$ nor $b=0$. Such elements are called **zero divisors.**

Definition. *A non-zero element a of a ring R is called a zero divisor or a divisor of zero if there exists an element* $b\neq 0\in R$ *such that either* $ab=0$ *or* $ba=0$. **(Jabalpur 1970)**

Rings without zero divisors. *A ring R is without zero-divisors if the product of no two non-zero elements of R is zero, i.e., if* $ab=0\Rightarrow a=0$ *or* $b=0$. **(Meerut 1984 P)**

On the other hand if in a ring R there exist non-zero elements a and b such that $ab=0$, then *R is said to be a ring with zero divisors.*

Example 1. *Suppose M is a ring of all* 2×2 *matrices with their elements as integers, the addition and multiplication of matrices being the two ring compositions. Then M is a ring with zero divisors.* **(Madras 1977)**

The null matrix $O=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$ is the zero element of this ring.

Now $A=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$, $B=\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}$ are two non-zero elements of this ring *i.e.*, $A\neq O$, $B\neq O$. We have

$$AB=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}=O.$$

Thus the product of two non-zero elements of the ring is equal to the zero element of the ring. Therefore M is a ring with zero divisors.

Also it is interesting to note that

$$BA=\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}=\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}\neq\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}.$$

Thus in a ring R it is possible that $ab=0$ but $ba\neq 0$.

Example 2. *The ring* $(\{0, 1, 2, 3, 4, 5\}, +_6, \times_6)$ *is a ring with zero divisors.* We have $2\times_6 3=0$, $3\times_6 4=0$ *i.e.*, the product of two non-zero elements is equal to the zero element of the ring.

Example 3. *The ring of integers is a ring without zero divisors.* The product of two non-zero integers cannot be equal to the zero integer.

Cancellation laws in a ring. If R is a ring then R is an abelian group with respect to addition. For addition composition the cancellation laws hold in all rings. Therefore the question of cancellation laws holding in a ring arises only for the multiplication composition.

We *say that cancellation laws hold in a ring R if*

$$a \neq 0,\ ab = ac \Rightarrow b = c$$

and $a \neq 0,\ ba = ca \Rightarrow b = c$ *where* $a, b, c \in R$. **(Meerut 1975)**

Theorem. *A ring R is without zero divisors if and only if the cancellation laws hold in R, i.e.,*

R is without zero divisors $\Leftrightarrow$ *cancellation laws hold in R.*

(I.A.S. 1974 ; Jabalpur 70 ; Madras 74 ; Garhwal 76; Kumāyon 78 ; Karnatak 77 ; Marathwada 74 ; Meerut 75, 81 ; Kanpur 71)

Proof. First suppose that R has no zero divisors. Let a, b, c be any three elements of R such that $a \neq 0$, $ab = ac$.

We have $ab = ac \Rightarrow ab - ac = 0 \Rightarrow a\,(b - c) = 0$.

Since R is without zero divisors, therefore

$$a\,(b - c) = 0 \text{ and } a \neq 0 \Rightarrow b - c = 0 \text{ i.e., } b = c.$$

Thus the left cancellation law holds in R. Similarly we can show that the right cancellation law also holds in R.

Conversely suppose that the cancellation laws hold in R. If possible let $ab = 0$, $a \neq 0$, $b \neq 0$.

Then we have $ab = a0$, since $a0 = 0$.

Now $a \neq 0$, $ab = a0 \Rightarrow b = 0$ by left cancellation law.

Thus we get a contradiction. Hence R is without zero divisors.

§ 6. Integral Domains. Fields. Division Rings.

Definition. Integral Domain. **(Madras 1973 ; Kolhapur 73 ; Garhwal 76; Kanpur 79; I.A.S. 71; Meerut 81, 84; Vikram 78; Andhra 77 ; Nagarjuna 79, 80 ; Madras 77)**

A ring is called an integral domain if it (i) is commutative, (ii) has unit element, (iii) is without zero divisors.

However some authors define an integral domain as a commutative ring without zero divisors. They do not require that an integral domain should definitely possess the unit element.

The most important example of an integral domain is the ring **I** of integers. We have proved that **I** is a commutative ring with unity. Also **I** does not possess zero divisors. We know that if a, b are integers such that $ab = 0$, then either a or b must be zero.

Also it can be seen that the algebraic structures $(\mathbf{C}, +, \cdot)$, $(\mathbf{Q}, +, \cdot)$, $(\mathbf{R}, +, \cdot)$, are all integral domains. As an example of a finite integral domain we have the ring $(\{0, 1, 2, 3, 4\}, +_5, \times_5)$.

Inversible elements in a ring with unity. In a ring every element possesses additive inverse. Therefore the question of an element being inversible or not arises only with respect to multiplication. *If R is a ring with unity, then an element $a \in R$ is called inversible, if there exists $b \in R$ such that $ab = 1 = ba$. Also then we write $b = a^{-1}$.*

Examples.

(i) 1 and -1 are the only two inversible elements of the ring of all integers.

(ii) $n \times n$ non-singular matrices with real numbers as elements are the only inversible elements of the ring of all $n \times n$ matrices with elements as real numbers.

Field. Definition. **(I.A.S. 1971 ; Kolhapur 73 ; Garhwal 76 ; Bombay 70 ; Nagarjuna 79, 80; Utkal 70 ; Andhra 77 ; Meerut 70, 81).**

A ring R with at least two elements is called a field if it,

(i) is commutative, (ii) has unity, (iii) is such that each non-zero element possesses multiplicative inverse.

For example, the ring of rational numbers $(Q, +, \cdot)$ is a field since it is a commutative ring with unity and each non-zero element is inversible.

The rings of real numbers and complex numbers are also examples of fields.

As an example of a finite field we have the ring

$$(\{0, 1, 2, 3, 4\}, +_5, \times_5).$$

If a, $0 \neq b$ are elements of a field F, then we shall often write $ab^{-1} = \frac{a}{b} = b^{-1}a$. In a field F, we have

$$\frac{a}{b} + \frac{c}{d} = (ab^{-1}) + (cd^{-1}) = (bd)^{-1} (bd) [(ab^{-1}) + (cd^{-1})]$$

$$= (bd)^{-1} [(bd)(ab^{-1}) + (bd)(cd^{-1})] = (bd)^{-1} (ad + bc) = \frac{ad+bc}{bd}.$$

[Note that multiplication in F is commutative].

Also $\frac{a}{b} \frac{c}{d} = (ab^{-1}) (cd^{-1}) = (ac) (b^{-1} d^{-1}) = (ac) (bd)^{-1} = \frac{ac}{bd}$.

Division ring or skew field. Definition.

(Andhra 1977; Allahabad 80; Meerut 81, 84 P)

A ring R with at least two elements is called a division ring or a skew field if it (i) has unity, (ii) is such that each non-zero element possesses multiplicative inverse.

Thus *a commutative division ring is a field.*

Every field is also a division ring. But a division ring is a field if it is also commutative. We shall later on give an example of a skew field which is not commutative *i.e.*, which is not a field.

Theorem 1. *Every field is an integral domain.*

(Delhi 1970; Garhwal 76; Jabalpur 70; Jiwaji 78; Nagarjuna 79, 80; Meerut 77, 78, 80, 81, 82)

Proof. Since a field F is a commutative ring with unity, therefore in order to show that every field is an integral domain we should show that a field has no zero divisors.

Let a, b be elements of F with $a \neq 0$ such that $ab = 0$.

Since $a \neq 0$, a^{-1} exists and we have

$$\begin{aligned} ab = 0 &\Rightarrow a^{-1}(ab) = a^{-1}\,0 \\ &\Rightarrow (a^{-1}a)\,b = 0 \\ &\Rightarrow 1b = 0 && [\because\ a^{-1}a = 1] \\ &\Rightarrow b = 0. && [\because\ 1b = b] \end{aligned}$$

Similarly, let $ab = 0$ and $b \neq 0$.

Since $b \neq 0$, b^{-1} exists and we have

$$\begin{aligned} ab = 0 &\Rightarrow (ab)\,b^{-1} = 0\,b^{-1} \\ &\Rightarrow a\,(bb^{-1}) = 0 \Rightarrow a1 = 0 \Rightarrow a = 0. \end{aligned}$$

Thus in a field $ab = 0 \Rightarrow a = 0$ or $b = 0$. Therefore a field has no zero divisors. Therefore every field is an integral domain.

But the converse is not true i.e., every integral domain is not a field. For example the ring of integers is an integral domain and it is not a field. The only inversible elements of the ring of integers are 1 and -1. **(Nagarjuna 1979, 80; Rajasthan 77; Meerut 81, 82)**

Note. For a field unity and zero are distinct elements *i.e.*, $1 \neq 0$. Let a be any non-zero element of a field. Then a^{-1} exists and is also non-zero. For, $a^{-1} = 0 \Rightarrow aa^{-1} = a0 \Rightarrow 1 = 0 \Rightarrow a1 = a0 \Rightarrow a = 0$ which is a contradiction. Now a field has no zero divisors. Therefore $1 = a^{-1}\,a \neq 0$.

Important. A field has no zero divisors. Therefore in a field the product of two non-zero elements will again be a non-zero element. Also the unit element $1 \neq 0$ and each non-zero element possesses multiplicative inverse which is again a non-zero element. The multiplication is commutative as well as associative. Therefore *the non-zero elements of a field form an abelian group with respect to multiplication.*

Theorem 2. *A sfield (skew field) has no divisors of zero.*

(Allahabad 1980)

Proof. Let D be a skew-field. Then D is a ring with unit element 1 and each non-zero element of D possesses multiplicative inverse.

Let a, b be elements of D with $a \neq 0$ such that $ab=0$.

Since $a \neq 0$, a^{-1} exists and we have

$$\begin{aligned} ab=0 &\Rightarrow a^{-1}(ab)=a^{-1}\,0 \\ &\Rightarrow (a^{-1}a)\,b=0 \Rightarrow 1b=0 \Rightarrow b=0. \end{aligned}$$

Similarly, let $ab=0$ with $b \neq 0$.

Since $b \neq 0$, b^{-1} exists and we have

$$\begin{aligned} ab=0 &\Rightarrow (ab)\,b^{-1}=0b^{-1} \\ &\Rightarrow a\,(bb^{-1})=0 \Rightarrow a1=0 \Rightarrow a=0. \end{aligned}$$

Therefore a skew field has no zero divisors.

Theorem 3. *A finite commutative ring without zero divisors is a field.*

Or

Every finite integral domain is a field.

(Delhi 1970; Nagarjuna 78; Gurunanak 75; Madras 78; Rajasthan 76; Kanpur 71; Garhwal 76; Poona 73; Banaras 70; Meerut 79, 80, 83, 84 P)

Proof. Let D be a finite commutative ring without zero divisors having n elements $a_1, a_2, \ldots, a_n$. In order to prove that D is a field, we must produce an element $1 \in D$ such that $1a=a \; \forall \; a \in D$. Also we should show that for every element $a \neq 0 \in D$ there exists an element $b \in D$ such that $ba=1$.

Let $a \neq 0 \in D$. Consider the n products $aa_1, aa_2, \ldots\, aa_n$.

All these are elements of D. Also they are all distinct. For suppose that $aa_i=aa_j$ for $i \neq j$.

Then $\qquad a\,(a_i-a_j)=0. \qquad \ldots(1)$

Since D is without zero divisors and $a \neq 0$, therefore (1) implies

$$a_i-a_j=0 \Rightarrow a_i=a_j, \text{ contradicting } i \neq j.$$

$\therefore$ $aa_1, aa_2, \ldots, aa_n$ are all the n distinct elements of D placed in some order. So one of these elements will be equal to a. Thus there exists an element, say, $1 \in D$ such that

$$a1=a=1a. \qquad [\because D \text{ is commutative}]$$

We shall show that this element 1 is the multiplicative identity of D. Let y be any element of D. Then from the above discussion for some $x \in D$, we shall have $ax=y=xa$.

$$\begin{aligned} \text{Now} \quad 1y &= 1\,(ax) && [\because ax=y] \\ &=(1a)\,x \\ &=ax && [\because 1a=a] \\ &=y && [\because ax=y] \end{aligned}$$

$=y1$ [$\because$ D is commutative]

Thus $1y=y=y1$, $\forall\ y \in D$. Therefore 1 is the unit element of the ring D.

Now $1 \in D$. Therefore from the above discussion one of the n products $aa_1, aa_2, \ldots, aa_n$ will be equal to 1. Thus there exists an element, say $b \in D$ such that

$$ab=1=ba.$$

$\therefore$ b is the multiplicative inverse of the non-zero element $a \in D$. Thus every non-zero element of D is inversible.

Hence D is a field.

Solved Examples

Ex. 1. *If a, b, c, d are elements of a ring R, then evaluate* $(a+b)(c+d)$.

Solution. We have

$(a+b)(c+d)=a(c+d)+b(c+d)$ [by right distributive law]

$=ac+ad+bc+bd$ [by left distributive law]

Ex. 2. *Prove that if $a, b \in R$ then $(a+b)^2=a^2+ab+ba+b^2$, where by x^2 we mean xx.*

Solution. We have

$(a+b)^2=(a+b)(a+b)$

$=a(a+b)+b(a+b)$ [by right distributive law]

$=(aa+ab)+(ba+bb)$ [by left distributive law]

$=a^2+ab+ba+b^2$.

Ex. 3. *If a, b are any elements of a ring R, prove that*

(*i*) $-(-a)=a$, (*ii*) $-(a+b)=-a-b$,

(*iii*) $-(a-b)=-a+b$.

Solution. (*i*) Since R is a group with respect to addition, therefore $-(-a)=a$. [Remember that in a group $(a^{-1})^{-1}=a$].

(*ii*) Since R is a group with respect to addition, therefore $-(a+b)=(-b)+(-a)$. But addition in R is commutative.

$\therefore$ $(-b)+(-a)=(-a)+(-b)=-a-b$.

$\therefore$ $-(a+b)=-a-b$.

(*iii*) $-(a-b)=-[a+(-b)]=-a+[-(-b)]=-a+b$.

Ex. 4. *If a, b, c, d are any elements of a ring R, prove that* $(a-b)(c-d)=(ac+bd)-(ad+bc)$.

Solution. We have

$(a-b)(c-d)=(a-b)c-(a-b)d$ [$\because$ $a(b-c)=ab-ac$]

$=(ac-bc)-(ad-bd)=(ac-bc)-ad+bd$

$=(ac+bd)-bc-ad$

[∵ addition is commutative and associative]

$=(ac+bd)-(bc+ad)$.

Ex. 5. *If R is a system satisfying all the conditions for a ring with unit element with the possible exception of $a+b=b+a$, prove that the axiom $a+b=b+a$ must hold in R and that R is thus a ring.*

(Kerala 1970; Delhi 70)

Solution. Since 1 is an element of R, we have

$(a+b)(1+1)=a(1+1)+b(1+1)$ [by right distributive law]

$=(a1+a1)+(b1+b1)=(a+a)+(b+b)$ ...(i)

Also $(a+b)(1+1)=(a+b)1+(a+b)1$

[by left distributive law]

$=(a+b)+(a+b)$ ...(ii)

[∵ 1 is the unit element]

From (i) and (ii), we get

$(a+a)+(b+b)=(a+b)+(a+b)$

$\Rightarrow [(a+a)+b]+b=[(a+b)+a]+b$ [by associativity of addition]

$\Rightarrow (a+a)+b=(a+b)+a$ [by right cancellation law for addition in R since with the given postulates R is a group with respect to addition].

$\Rightarrow a+(a+b)=a+(b+a)$ [by associativity of addition in R]

$\Rightarrow (a+b)=(b+a)$ [by left cancellation law for addition in R]

Thus addition is commutative in R. Hence R is a ring.

Ex. 6. *If R is a ring such that $a^2=a\ \forall\ a \in R$ prove that*

(i) $a+a=0\ \forall\ a \in R$ *i.e., each element of R is its own additive inverse.* **(Nagpur 1970)**

(ii) $a+b=0 \Rightarrow a=b$. *(iii) R is a commutative ring.*

(Madras 1974; Banaras 70; Nagarjuna 79, 80)

Solution. (i) $a \in R \Rightarrow a+a \in R$.

Now $(a+a)^2=(a+a)$ [given]

$\Rightarrow (a+a)(a+a)=a+a$

$\Rightarrow (a+a)a+(a+a)a=a+a$ [Left Dist. Law]

$\Rightarrow (a^2+a^2)+(a^2+a^2)=a+a$ [Right Dist. Law]

$\Rightarrow (a+a)+(a+a)=a+a$ [∵ $a^2=a$]

$\Rightarrow (a+a)+(a+a)=(a+a)+0$ [∵ $a+0=a$]

$\Rightarrow (a+a)=0$ [by left cancellation law for addition in R]

(ii) We have just proved that $a+a=0$.

∴ $a+b=0 \Rightarrow a+b=a+a \Rightarrow b=a$, by left cancellation law for addition in R.

(iii) We have

$$\begin{aligned}
& (a+b)^2=(a+b) \\
\Rightarrow\ & (a+b)\,(a+b)=(a+b) \\
\Rightarrow\ & (a+b)\,a+(a+b)\,b=a+b && \text{[Left Dist. Law]} \\
\Rightarrow\ & (a^2+ba)+(ab+b^2)=a+b && \text{[Right Dist. Law]} \\
\Rightarrow\ & (a+ba)+(ab+b)=a+b && [\because\ a^2=a,\ b^2=b] \\
\Rightarrow\ & (a+b)+(ba+ab)=(a+b)+0 && \text{[by commutativity and associativity of addition]} \\
\Rightarrow\ & ba+ab=0 && \text{[by left cancellation law for addition in } R] \\
\Rightarrow\ & ab=ba. && \text{[by part (ii) of this question]}
\end{aligned}$$

$\therefore$ R is commutative ring.

Note. An element a of a ring R is said to be idempotent if $a^2=a$. *A ring R is called a* **Boolean Ring** *if all of its elements are idempotent i.e., if $a^2=a\ \forall\ a\in R$.* (Nagarjuna 1979, 80)

Ex. 7. *Prove that the set M of 2×2 matrices over the field of real numbers is a ring with respect to matrix addition and multiplication. Is it a commutative ring with unity element ? Find the zero element. Does this ring possess zero divisors ?*

(Delhi 1970; Madras 77; Mysore 77)

Solution. Let $A, B\in M$. Then $A+B\in M$ and $AB\in M$. Therefore M is closed with respect to addition and multiplication of matrices.

Both addition and multiplication of matrices are associative compositions.

$\therefore$ $A+(B+C)=(A+B)+C\ \forall\ A, B, C\in M$

and $A(BC)=(AB)\,C\ \forall\ A, B, C\in M.$

Addition of matrices is a commutative composition. Therefore for all $A, B\in M$, we have $A+B=B+A$.

If O be the null matrix of the type 2×2, then $O\in M$ and $O+A=A\ \forall\ A\in M$.

Further multiplication of matrices is distributive with respect to addition.

$\therefore$ $A\,(B+C)=AB+AC$

and $(B+C)\,A=BA+CA\ \forall\ A, B, C\in M.$

$\therefore$ M is a ring with respect to the given compositions.

Multiplication of matrices is not in general a commutative composition. For example, if $A=\begin{bmatrix}2 & 4\\ 3 & 5\end{bmatrix}$, $B=\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}$, then

$AB=\begin{bmatrix}2 & 4\\ 3 & 5\end{bmatrix}\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}=\begin{bmatrix}2 & 8\\ 3 & 11\end{bmatrix}$ and $BA=\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}\begin{bmatrix}2 & 4\\ 3 & 5\end{bmatrix}=\begin{bmatrix}8 & 14\\ 3 & 5\end{bmatrix}$.

Thus $AB \neq BA$ and so the ring is a non-commutative ring.

If I be the unit matrix of the type 2×2 *i.e.*, if $I=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$, then $I \in M$. Also we have $AI=A=IA \;\forall\; A\in M$.

$\therefore$ I is the multiplicative identity.

Thus the ring possesses the unit element and we have $I=1$ (the unit element of the ring).

The null matrix $O=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$ is the additive identity and is therefore the zero element of the ring *i.e.*, $O=0$ (the zero element of the ring).

The ring possesses zero divisors. For example if

$A=\begin{bmatrix}0 & 1\\ 0 & 1\end{bmatrix}$, $B=\begin{bmatrix}2 & 3\\ 0 & 0\end{bmatrix}$, then $AB=\begin{bmatrix}0 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}2 & 3\\ 0 & 0\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$.

Thus the product of two non-zero elements of the ring is equal to the zero element of the ring.

Ex. 8. *Do the following sets form integral domains with respect to ordinary addition and multiplication ? If so state if they are fields.*

(*i*) *the set of numbers of the form $b\sqrt{2}$ with b rational.*

(*ii*) *the set of even integers.* **(Meerut 1975)**

(*iii*) *the set of positive integers.* **(Meerut 1968)**

Solution. (i) Let $A=\{b\sqrt{2} : b\in \mathbf{Q}\}$.

We have $3\sqrt{2}\in A$ and $5\sqrt{2}\in A$. Then $(3\sqrt{2})(5\sqrt{2})=30$.

Now 30 cannot be put in the form $b\sqrt{2}$ where b is a rational number. Therefore $30 \notin A$. Thus A is not closed with respect to multiplication. Therefore the question of A becoming a ring does not arise.

(ii) Let R be the set of all even integers. Then R is a ring with respect to addition and multiplication of integers. Also the multiplication is a commutative composition. R is without zero divisors since the product of two non-zero even integers cannot be equal to zero which is the zero element of this ring. Since the integer $1 \notin R$, therefore R is a ring without unity.

R will be an integral domain if we do not require the existence of the unit element for an integral domain.

But R is not a field since the multiplicative identity does not exist.

(iii) Let N be the set of positive integers. Since the integer $0 \notin N$, therefore the additive identity does not exist. So N will not be a ring.

Ex 9. *Show that the set of numbers of the form* $a+b\sqrt{2}$, *with a and b as rational numbers is a field.*

(Kumayon 1977 ; Kanpur 80 ; Agra 72 ; Meerut 73)

Solution. Let $R=\{a+b\sqrt{2} : a, b \in \mathbf{Q}\}$.

Let $a_1+b_1\sqrt{2} \in R$ and $a_2+b_2\sqrt{2} \in R$. Then

$$a_1, b_1, a_2, b_2 \in \mathbf{Q}.$$

We have $(a_1+b_1\sqrt{2})+(a_2+b_2\sqrt{2})=(a_1+a_2)+(b_1+b_2)\sqrt{2}$

$\in R$ since a_1+a_2, $b_1+b_2 \in \mathbf{Q}$.

Also $(a_1+b_1\sqrt{2})\,(a_2+b_2\sqrt{2})=(a_1a_2+2b_1b_2)+(a_1b_2+a_2b_1)\sqrt{2}$

$\in R$ since $a_1a_2+2b_1b_2$, $a_1b_2+a_2b_1 \in \mathbf{Q}$.

Thus R is **closed with respect to addition and multiplication.**

All the elements of R are real numbers and we know that addition and multiplication are both associative as well as commutative compositions in the set of real numbers.

Further we have $0+0\sqrt{2} \in R$ since $0 \in \mathbf{Q}$.

If $a+b\sqrt{2} \in R$, then

$(0+0\sqrt{2})+(a+b\sqrt{2})=(0+a)+(0+b)\sqrt{2}=a+b\sqrt{2}$.

$\therefore$ $0+0\sqrt{2}$ is the **additive identity.**

Again if $a+b\sqrt{2}\in R$, then $(-a)+(-b)\sqrt{2} \in R$ and we have $[(-a)+(-b)\sqrt{2}]+(a+b\sqrt{2})=0+0\sqrt{2}$.

$\therefore$ each element of R possesses **additive inverse.**

Further in the set of real numbers **multiplication is distributive with respect to addition.**

Again $1+0\sqrt{2} \in R$ and we have

$(1+0\sqrt{2})\,(a+b\sqrt{2})=a+b\sqrt{2}=(a+b\sqrt{2})\,(1+0\sqrt{2})$.

$\therefore$ $1+0\sqrt{2}$ is the **multiplicative identity.**

Thus R is a commutative ring with unity. The zero element of the ring is $0+0\sqrt{2}$ and the unit element is $1+0\sqrt{2}$.

Now R will be a field if each non-zero element of R possesses multiplicative inverse.

Let $a+b\sqrt{2}$ be any non-zero element of this ring *i.e.*, at least one of a and b is not zero.

Then $\dfrac{1}{a+b\sqrt{2}}=\dfrac{a-b\sqrt{2}}{(a+b\sqrt{2})\,(a-b\sqrt{2})}=\dfrac{a-b\sqrt{2}}{a^2-2b^2}$

$$=\left(\frac{a}{a^2-2b^2}\right)+\left(-\frac{b}{a^2-2b^2}\right)\sqrt{2}.$$

Now if a and b are rational numbers, then we can have $a^2=2b^2$ only if $a=0$, $b=0$. Since here at least one of the rational numbers a and b is not 0, therefore we cannot have $a^2=2b^2$ *i.e.*, $a^2-2b^2=0$.

∴ $\frac{a}{a^2-2b^2}$ and $-\frac{b}{a^2-2b^2}$ are both rational numbers and at least one of them is not zero.

∴ $\left(\frac{a}{a^2-2b^2}\right)+\left(-\frac{b}{a^2-2b^2}\right)\sqrt{2}$ is a non-zero element of R and is the multiplicative inverse of $a+b\sqrt{2}$.

Hence the given system is a field.

Ex. 10. *Prove that the set $I(\sqrt{2})$ of all real numbers of the form $a+b\sqrt{2}$, with a and b as integers is an integral domain with respect to ordinary addition and multiplication. Is it a field ?*

(Kanpur 1970; Agra 74; Rajasthan 77; Meerut 81; Bombay 70)

Solution. As in Ex. 9, we can easily verify that the given system is a commutative ring with unity element $1+0\sqrt{2}$. Also $0+0\sqrt{2}$ is the zero element of this ring. Now in order to prove that this ring is an integral domain, we should prove that this ring is without zero divisors.

Let $a+b\sqrt{2}$ and $c+d\sqrt{2}$ be any two elements of this ring. Then

$$(a+b\sqrt{2})(c+d\sqrt{2})=0+0\sqrt{2}$$

$\Rightarrow ac+2bd=0$ and $bc+ad=0$ and this will happen only when either

$$a=0 \text{ and } b=0$$

or

$$c=0 \text{ and } d=0.$$

Thus $(a+b\sqrt{2})(c+d\sqrt{2})=0+0\sqrt{2}$

$\Rightarrow$ either $a+b\sqrt{2}=0$ or $c+d\sqrt{2}=0$.

Thus the given ring is without zero divisors. Therefore it is an integral domain. But it is not a field. Obviously $5+3\sqrt{2}$ is a non-zero element of this ring. Its inverse would have been

$$\frac{1}{5+3\sqrt{2}}=\frac{5-3\sqrt{2}}{(5+3\sqrt{2})(5-3\sqrt{2})}=\frac{5-3\sqrt{2}}{7}=\frac{5}{7}-\frac{3}{7}\sqrt{2}$$

which is not an element of this ring.

Ex. 11. *A* **Gaussian integer** *is a complex number $a+ib$, where a and b are integers. Show that the set $J[i]$ of Gaussian integers* *forms a ring under ordinary addition and multiplication of complex numbers. Is it an integral domain? Is it a field?*

(I.A.S. 1973; Kolhapur 74; Poona 74; Madurai 78)

Solution. Let $a+ib$ and $c+id$ be any two Gaussian integers. Then

$$(a+ib)+(c+id)=(a+c)+i\,(b+d)$$

and

$$(a+ib)\,(c+id)=(ac-bd)+i\,(ad+bc).$$

These are again Gaussian integers. Therefore $J[i]$ is closed with respect to ordinary addition and multiplication of complex numbers.

Further in complex numbers both addition and multiplication are associative as well as commutative compositions. Also multiplication distributes with respect to addition. The Gaussian integer $0+i0$ is the additive identity. The additive inverse of $a+ib$ is $(-a)+i\,(-b)$. The Gaussian integer $1+i0$ is the multiplicative identity.

Therefore the set of Gaussian integers is a commutative ring with unity for the given composition.

Also this ring is free from zero divisors since the product of two non-zero complex numbers cannot be zero. Therefore $J\,[i]$ is an integral domain.

But this is not a field since the multiplicative inverse of $a+ib$ will be $\dfrac{a}{a^2+b^2}+i\left(-\dfrac{b}{a^2+b^2}\right)$ which is not always a Gaussian integer as $\dfrac{a}{a^2+b^2}$ and $-\dfrac{b}{a^2+b^2}$ are not necessarily integers.

Ex. 12. *Prove that the totality R of all ordered pairs (a, b) of real numbers is a commutative ring with zero divisors under the addition and multiplication of ordered pairs defined as*

$$(a,\,b)+(c,\,d)=(a+c,\,b+d)$$
$$(a,\,b)\,(c,\,d)=(ac,\,bd)$$
$$\forall\ (a,\,b),\,(c,\,d)\in R.$$

(Sambalpur 1977)

Solution. We see that R is closed with respect to the two compositions since $a+c$, $b+d$, ac, bd are all real numbers. Now let $(a,\,b)$, $(c,\,d)$, $(e,\,f)$ be any elements of R. Then we observe :

Associativity of addition. We have

$$\begin{aligned}[(a,\,b)+(c,\,d)]+(e,\,f)&=(a+c,\,b+d)+(e,\,f)\\&=([a+c]+e,\,[b+d]+f)\\&=(a+[c+e],\,b+[d+f])\\&\qquad[\because\ \text{addition of real numbers is associative}]\\&=(a,\,b)+(c+e,\,d+f)\end{aligned}$$

$$=(a, b)+[(c, d)+(e, f)].$$

$\therefore$ addition in R is associative.

Commutativity of addition. We have

$$(a, b)+(c, d)=(a+c, b+d)=(c+a, d+b)=(c, d)+(a, b).$$

Existence of additive identity. We have $(0, 0) \in R$. Also $(0, 0)+(a, b)=(0+a, 0+b)=(a, b)$.

Existence of additive inverse. If $(a, b)\in R$, then $(-a, -b)\in R$ and we have $(-a, -b)+(a, b)=(-a+a, -b+b)=(0, 0)$.

$\therefore$ $(-a, -b)$ is the additive inverse of (a, b).

Associativity of Multiplication. We have

$$[(a, b)\,(c, d)]\,(e, f)=(ac, bd)\,(e, f)=([ac]\,e, [bd]\,f)$$
$$=(a\,[ce], b\,[df]).$$

[$\because$ multiplication of real numbers is associative]

$$=(a, b)\,(ce, df)=(a, b)\,[(c, d)\,(e, f)].$$

Distributive laws. We have

$$(a, b)\,[(c, d)+(e, f)]=(a, b)\,(c+e, d+f)$$
$$=(a\,[c+e], b\,[d+f])$$
$$=(ac+ae, bd+bf)=(ac, bd)+(ae, bf)$$
$$=(a, b)\,(c, d)+(a, b)\,(e, f).$$

Similarly we can show that the other distributive law also olds good.

$\therefore$ R is a ring with respect to the given compositions.

Commutativity of multiplication. We have

$$(a, b)\,(c, d)=(ac, bd)=(ca, db)=(c, d)\,(a, b).$$

$\therefore$ R is a commutative ring.

Existence of multiplicative identity. We have $(1, 1) \in R$. If $(a, b) \in R$, then $(1, 1)\,(a, b)=(1a, 1b)=(a, b)=(a, b)\,(1, 1)$.

$\therefore$ $(1, 1)$ is the multiplicative identity and is therefore the unit element of the ring. So R is a ring with unity also.

The zero element of this ring is the ordered pair $(0, 0)$.

Now in order to show that R is a ring with zero divisors we should show that there exist two non-zero elements of R whose product is equal to the zero element of R. Obviously neither $(3, 0)$ nor $(0, 5)$ is equal to the zero element of R. But $(3, 0)\,(0, 5)=(3\times 0, 0\times 5=(0, 0)$ which is the zero element of R.

$\therefore$ R is a ring with zero divisors.

Ex. 13. *Let* **C** *be the set of the ordered pairs* (a, b) *of real numbers. Define addition and multiplication in* **C** *by the equations*

$$(a, b)+(c, d)=(a+c, b+d)$$
$$(a, b)\,(c, d)=(ac-bd, bc+ad).$$

Prove that **C** *is a field.* **(Meerut 1978; Marathwada 74)**

Proof. We see that C is closed with respect to the two compositions since $a+c, b+d, ac-bd, bc+ad$ are all real numbers. Now let $(a, b), (c, d), (e, f)$ be any elements of C. Then we make the following observations :

Associativity of addition. We have

$$[(a, b)+(c, d)]+(e, f)=(a+c, b+d)+(e, f)$$
$$=([a+c]+e, [b+d]+f)=(a+[c+e], b+[d+f])$$
$$=(a, b)+(c+e, d+f)=(a, b)+[(c, d)+(e, f)].$$

Commutativity of addition. We have

$$(a, b)+(c, d)=(a+c, b+d)=(c+a, d+b)=(c, d)+(a, b).$$

Existence of additive Identity. We have $(0, 0)\in$ C.

Also $(0, 0)+(a, b)=(0+a, 0+b)=(a, b)$.

$\therefore$ (0, 0) is the additive identity.

Existence of additive inverse. If $(a, b) \in$ C, then $(-a, -b)\in$ C. We have

$$(-a, -b)+(a, b)=(-a+a, -b+b)=(0, 0)$$

$\therefore$ $(-a, -b)$ is the additive inverse of (a, b).

Associativity of multiplication. We have

$$[(a, b)(c, d)](e, f)=(ac-bd, bc+ad)(e, f)$$
$$=([ac-bd]e-[bc+ad]f, [bc+ad]e+[ac-bd]f)$$
$$=(a[ce-df]-b[de+cf], b[ce-df]+a[de+cf])$$
$$=(a, b)(ce-df, de+cf)=(a, b)[(c, d)(e, f)].$$

Distributive laws We have

$$(a, b)[(c, d)+(e, f)]=(a, b)(c+e, d+f)$$
$$=(a[c+e]-b[d+f], b[c+e]+a[d+f])$$
$$=([ac-bd]+[ae-bf], [bc+ad]+[be+af])$$
$$=(ac-bd, bc+ad)+(ae-bf, be+af)=(a, b)(c, d)+(a, b)(e, f).$$

Similarly we can show that the other distributive law also holds good.

Therefore C is a ring with respect to the two compositions. The ordered pair (0, 0) is the zero element of the ring.

Commutativity of multiplication. We have

$$(a, b)(c, d)=(ac-bd, bc+ad)=(ca-db, da+cb)=(c, d)(a, b).$$

Existence of multiplicative identity. We have $(1, 0) \in$ C. If $(a, b)\in$ C, then $(a, b)(1, 0)=(a1-b0, b1+a0)=(a, b)=(1, 0)(a, b)$. Therefore (1, 0) is the unity element of the ring.

Existence of multiplicative inverse of each non-zero element of C. Let (a, b) be any non-zero element of C. Then a and b are not both simultaneously zero. If (c, d) is the multiplicative inverse of (a, b), then we should have

$$(a, b)\ (c, d)=(1, 0)$$

or
$$(ac-bd,\ bc+ad)=(1, 0).$$

By the definition of the equality of two ordered pairs, we have

$$ac-bd=1 \text{ and } bc+ad=0.$$

Solving these equations for c, d, we get

$$c=\frac{a}{a^2+b^2},\ d=\left(-\frac{b}{a^2+b^2}\right).$$

Now $a\neq 0$ or $b\neq 0 \Rightarrow a^2+b^2\neq 0$. Therefore either c or d or both are non-zero real numbers. Thus $\left(\frac{a}{a^2+b^2},\ -\frac{b}{a^2+b^2}\right)$ is the multiplicative inverse of (a, b). Hence **C** is a field.

Note. In this question **C** is nothing but the set of complex numbers defined as ordered pairs of real numbers. Thus we have proved that the set of complex numbers is a field with respect to addition and multiplication of complex numbers.

Ex. 14. *Show that the set R of all real valued continuous functions defined in the closed interval* [0, 1] *is a commutative ring with unity with respect to the addition and multiplication of functions defined* **pointwise** *as follows :*

$$(f+g)\ (x)=f\ (x)+g\ (x)$$

and
$$(fg)\ (x)=f\ (x)\ g\ (x),$$

where f, g are any two members of R. **(Andhra 1975)**

Solution. If f is a real valued function in the closed interval [0, 1], then we mean that $f\ (x)$ is a real number $\forall\ x \in [0, 1]$. We know that the sum and product of two real numbers are also real numbers. Also the sum and product of two continuous functions is also a continuous function. Therefore R is closed with respect to the given compositions.

Now let f, g, h be any three elements of R. Then we make the following observations.

Associativity of addition. For every $x \in [0, 1]$, we have

$$[(f+g)+h]\ (x)=[(f+g)\ (x)]+h\ (x)$$
$$=[f\ (x)+g\ (x)]+h\ (x)=f\ (x)+[g\ (x)+h\ (x)]$$

[$\because$ $f(x)$, $g\ (x)$, $h\ (x)$ are real numbers and addition of real numbers is associative]

$$=f\ (x)+(g+h)\ (x)=[f+(g+h)]\ (x).$$

$\therefore$ $(f+g)+h=f+(g+h)$, by the equality of two mappings.

Commutativity of addition. We have

$$(f+g)\ (x)=f\ (x)+g\ (x)=g\ (x)+f\ (x)$$

[$\because$ addition of real numbers is commutative]

$$=(g+f)(x).$$

$$\therefore\ f+g=g+f.$$

Existence of additive identity. Let us define a function e by the rule $e(x)=0\ \forall\ x\in[0,1]$. Then $e\in R$. Also if $f\in R$, we have $(e+f)(x)=e(x)+f(x)=0+f(x)=f(x)$.

$$\therefore\ e+f=f.$$

$\therefore$ the function e is the additive identity.

Existence of additive inverse. Let $f\in R$. Let us define a function $-f$ by the formula $(-f)(x)=-[f(x)]\ \forall\ x\in[0,1]$.
Then $-f\in R$ and we have

$$[-f+f](x)=(-f)(x)+f(x)=-f(x)+f(x)=0=e(x).$$

$$\therefore\ -f+f=e.$$

$\therefore$ the function $-f$ is the additive inverse of f.

Associativity of multiplication. We have

$$[(fg)h](x)=[(fg)(x)]\,h(x)$$
$$=[f(x)\,g(x)]\,h(x)=f(x)\,[g(x)\,h(x)]$$
$$=f(x)\,[(gh)(x)]=[f(gh)](x).$$
$$\therefore\ (fg)h=f(gh).$$

Distributive laws. We have

$$[f(g+h)](x)=f(x)\,[(g+h)(x)]=f(x)\,[g(x)+h(x)]$$
$$=f(x)\,g(x)+f(x)\,h(x)=(fg)(x)+(fh)(x)$$
$$=[fg+fh](x).$$

$\therefore\ f(g+h)=fg+fh.$

Similarly we can prove that $(g+h)f=gf+hf$.

$\therefore$ R is a ring with respect to the given compositions.

Commutativity of multiplication. We have

$$(fg)(x)=f(x)\,g(x)=g(x)\,f(x)=(gf)(x).$$

$\therefore\ fg=gf.$

$\therefore$ R is a commutative ring.

Existence of multiplicative identity. Let us define a function i by the formula $i(x)=1\ \forall\ x\in[0,1]$. Then $i\in R$. If $f\in R$, we have $(if)(x)=i(x)\,f(x)=1\,f(x)=f(x)$.

$\therefore\ if=f=fi$ [by commutativity of multiplication]

$\therefore$ The function i is the multiplicative identity.

Thus the ring R is with unity element.

Ex. 15. *Give an example of a skew field which is not a field.*

[Delhi 1970 ; Poona 73 ; Meerut 79, 83 (P)]

Solution. Let M be the set of all 2×2 matrices of the form

$$\begin{bmatrix} a+ib & c+id \\ -c+id & a-ib \end{bmatrix},$$

where a, b, c, d are arbitrary real numbers.

First we shall prove that M is a ring with respect to addition and multiplication of matrices.

Let $A=\begin{bmatrix} a+ib & c+id \\ -c+id & a-ib \end{bmatrix}$ and $B=\begin{bmatrix} p+iq & r+is \\ -r+is & p-iq \end{bmatrix}$ be any two elements of M. We have

$$A+B=\begin{bmatrix} (a+p)+i(b+q) & (c+r)+i(d+s) \\ -(c+r)+i(d+s) & (a+p)-i(b+q) \end{bmatrix},$$ which is obviously a matrix of the given form. So $A+B \in M$.

Also AB

$$\begin{bmatrix} (a+ib)(p+iq)+(c+id)(-r+is) & (a+ib)(r+is)+(c+id)(p-q) \\ (-c+id)(p+iq)+(a-ib)(-r+is) & (-c+id)(r+is)+(a-ib)(p-iq) \end{bmatrix}$$

$$\begin{bmatrix} (ap-bq-cr-ds) & (ar-bs+cp+dq) \\ \quad +i\,(aq+bp+cs-dr) & \quad +i(as+br-cq+dp) \\ -(cp+dq+ar-bs) & (-cr-ds+ap-bq) \\ \quad +i(dp-cq+as+br) & \quad -i(cs-dr+aq+bp) \end{bmatrix}$$

which is obviously an element of M. Thus M is closed with respect to addition and multiplication of matrices. Further matrix addition is commutative as well as associative. The zero matrix $\begin{bmatrix} 0+i0 & 0+i0 \\ -0+i0 & 0-i0 \end{bmatrix}$ is the additive identity and so it is the zero element of M. If $A=\begin{bmatrix} a+ib & c+id \\ -c+id & a-ib \end{bmatrix} \in M$, then obviously $-A=\begin{bmatrix} -a-ib & -c-id \\ c-id & -a+ib \end{bmatrix} \in M$. Thus each element of M possesses additive inverse. Further matrix multiplication is associative and it is distributive with respect to addition. Hence M is a ring with respect to addition and multiplication of matrices.

Existence of multiplicative identity. The matrix $\begin{bmatrix} 1+i0 & 0+0i \\ -0+i0 & 1-i0 \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is obviously an element of M. It is the unit matrix and so it is the multiplicative identity. Thus M is a ring with unity.

Existence of multiplicative inverse of each non-zero element of M. Let $A=\begin{bmatrix} a+ib & c+id \\ -c+id & a-ib \end{bmatrix}$ be any non-zero element of M *i.e.*, a, b, c, d are not all equal to zero. We have $|A|$ *i.e.*, det

$A=a^2+b^2+c^2+d^2\neq 0$. Therefore the matrix A is non-singular and is therefore inversible. We must show that $A^{-1}\in M$. Let $|A|=m$. Then m is a real number and $m>0$. We have

$$A^{-1}=\frac{1}{|A|}\text{ Adj } A=\frac{1}{m}\begin{bmatrix} a-ib & -c-id \\ c-id & a+ib \end{bmatrix},$$

which is obviously an element of M. Therefore each non-zero element of M possesses multiplicative inverse.

$\therefore$ M is a skew field (or a division ring).

Now we shall show that M is not a field *i.e.*, multiplication in M is not commutative. We have $A=\begin{bmatrix} 3+4i & 5+6i \\ -5+6i & 3-4i \end{bmatrix}\in M$,

and $B=\begin{bmatrix} 1+i0 & 1+i0 \\ -1+i0 & 1-i0 \end{bmatrix}=\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\in M$.

Also $AB=\begin{bmatrix} -2-2i & 8+10i \\ -8+10i & -2+2i \end{bmatrix}$, and $BA=\begin{bmatrix} -2+10i & 8+2i \\ -8+2i & -2-10i \end{bmatrix}$.

We see that $AB\neq BA$. Therefore multiplication in M is not commutative and so M is not a field.

Hence M is a skew field which is not a field.

Ex. 16. *Let p be a prime number. Prove that the set of integers* $\mathbf{I}_p$,

$$\mathbf{I}_p=\{0,\ 1,\ 2,\ 3,\ldots\ldots,\ p-1\}$$

forms a field with respect to addition and multiplication modulo p.

(Mysore 1970; Andhra 77;
Madras 77; Delhi 70; Utkal 70; Aligarh 67)

Solution. As in the chapter on groups we can prove that $(\mathbf{I}_p, +_p)$ is an abelian group. Also $\mathbf{I}_p$ is closed with respect to '$\times_p$'. The composition '$\times_p$' is associative as well as commutative. Also '$\times_p$' is distributive with respect to '$+_p$' as can be proved as below :

Let $a, b, c\in \mathbf{I}_p$. Then

$a\times_p(b+_p c)=a\times_p(b+c)$ $\quad[\because\ b+_p c\equiv b+c \pmod p)]$

$=$least non-negative remainder when $a\ (b+c)$ is divided by p

$=$least non-negative remainder when $ab+ac$ is divided by p

$=(ab)+_p(ac)$

$=(a\times_p b)+_p(ac)$ $\quad[\because\ ab\equiv a\times_p b \pmod p)]$

$=(a\times_p b)+_p(a\times_p c)$.

Similarly we can prove the other distributive law.

1 is the identity element for '$\times_p$'.

The zero element of the ring $(\mathbf{I}_p, +_p, \times_p)$ is 0. As in groups, we can prove that each non-zero element of I_p has multiplicative inverse.

In short the non-zero elements of I_p form an abelian group with respect to '$\times_p$'. Therefore $(I_p, +_p, \times_p)$ is a field.

Note. If p is not prime, then this ring will have zero divisors and so it cannot be a field.

Ex. 17. *Prove that the set of residue classes modulo p is a commutative ring with respect to addition and multiplication of residue classes. Further show that the ring of residue classes modulo p is a field if and only if p is a prime.*

(Kanpur 1980; Vikram 79; Karnatak 77; I A.S 74; Allahabad 80; Kumayun 77; Sagar 77; Bombay 70; Meerut 74)

Solution. Let I_p be the set of residue classes modulo p. Then I_p has p distinct elements. Thus $I_p = \{[0], [1], [2], \ldots, [p-1]\}$.

Let $[a], [b] \in I_p$. Then we define addition and multiplication of residue classes as follows :

$$[a]+[b]=[a+b], \qquad \text{(addition)}$$
$$[a]\,[b]=[ab] \qquad \text{(multiplication).}$$

Since $[a+b]$ and $[ab]$ are both residue classes modulo p, therefore I_p is closed with respect to addition and multiplication

Now let $[a], [b], [c]$ be any elements of I_p. Then we observe :

Commutativity of addition.

$$[a]+[b]=[a+b] \qquad \text{(by def. of addition of residue classes)}$$
$$=[b+a] \qquad (\because \text{ addition of integers is commutative})$$
$$=[b]+[a].$$

Associativity of addition. We have

$$([a]+[b])+[c]=[a+b]+[c]=[(a+b)+c]=[a+(b+c)]$$
$$=[a]+[b+c]=[a]+([b]+[c]).$$

Existence of additive identity. We have $[0] \in I_p$. If $[a] \in I_p$, then $[0]+[a]=[0+a]=[a]$. Therefore $[0]$ is the additive identity.

Existence of additive inverse. Let $[a] \in I_p$. Then $[-a] \in I_p$. We have $[-a]+[a]=[-a+a]=[0]$. Therefore $[-a]$ is the additive inverse of $[a]$.

Associativity of multiplication. We have

$$([a]\,[b])\,[c]=[ab]\,[c]=[(ab)\,c]=[a\,(bc)]=[a]\,[bc]=[a]\,([b]\,[c])$$

Commutativity of multiplication. We have

$$[a]\,[b]=[ab]=[ba]=[b]\,[a].$$

Distributive Laws. We have

$$[a]\,([b]+[c])=[a]\,[b+c]=[a\,(b+c)]=[ab+ac]$$

$=[ab]+[ac]=[a]\,[b]+[a]\,[c]$.

Similarly $([b]+[c])\,[a]=[b]\,[a]+[c]\,[a]$.

Thus $\mathbf{I}_p$ is a commutative ring. Note that it is a finite ring because it has p elements. Now suppose that p is a prime number. Then to prove that $\mathbf{I}_p$ is a field. Let $[a], [b] \in \mathbf{I}_p$. Then

$[a]\,[b]=[0] \Rightarrow [ab]=[0]$

$\Rightarrow p$ is a divisor of ab *i.e.*, $p \mid ab$

$\Rightarrow p \mid a$ or $p \mid b$. [Note that if a and b are any two integers and p is a prime number, then $p \mid ab \Rightarrow p \mid a$ or $p \mid b$]

$\Rightarrow [a]=[0]$ or $[b]=[0]$.

Thus $\mathbf{I}_p$ is without zero divisors. Therefore $\mathbf{I}_p$ is an integral domain. But every finite integral domain is a field. Hence $\mathbf{I}_p$ is a field.

Conversely suppose that $\mathbf{I}_p$ is a field. Then $\mathbf{I}_p$ is an integral domain. Therefore $\mathbf{I}_p$ is without zero divisors. We are to prove that p is a prime number. Suppose p is not prime, but p is composite. Let $p=mn$, where $1 < m < p$, $1 < n < p$. Then

$$[mn]=[p]$$

$$\Rightarrow [m]\,[n]=[0] \qquad (\because \; [p]=[0])$$

Also $[m]\neq[0]$, since $1 < m < p$. Similarly $[n]\neq 0$.

Thus $[m]\,[n]=[0]$ while neither $[m]=[0]$ nor $[n]=[0]$.

Therefore $\mathbf{I}_p$ possesses zero divisors and thus we get a contradiction. Hence p must be prime.

Exercises

1. Prove that the set of rational numbers (real numbers or complex numbers) is a field with respect to addition and multiplication. **(Meerut 1976)**

2. (i) Prove that the set of integers is an integral domain with respect to addition and multiplication. **(Meerut 1975)**

(ii) Define a field. Prove that every field is an integral domain, but there exist some integral domains which are not fields. **(Madurai 1978)**

3. Define a ring and furnish an example of (i) a non-commutative ring with unity, (ii) a commutative ring without unity. **(Kerala 1974 ; Delhi 70)**

4. Prove that in the list of axioms for a ring R with unity the axiom demanding commutativity under addition may be omitted. **(Kerala 1970)**

5. Is every field also a division ring ? Does the set of all integers under usual addition and multiplication form a field ? Give some example of a field which is finite.

(Kanpur 1969 ; Meerut 76)

6. Prove that the set of integers R,

$$R=\{0, 1, 2, 3, 4\}$$

forms a field under addition modulo 5 and multiplication modulo 5.

7. Prove that the set $\{0, 1, 2\}$ (mod 3) is a field with respect to addition and multiplication. **(Meerut 1977)**

8. If two operations $*$ and o on the set **I** of integers are defined as follows :

$$a * b=a+b \quad 1,\ a\ o\ b=a+b-ab$$

prove that the system $(\mathbf{I}, *, o)$ is a commutative ring with identity.

(Banaras 1970)

9. If R is a ring with unity element 1, then $(-1)\,a=-a=a\,(-1)\ \forall\ a \in R$ and $(-1)\,(-1)=1$. **(Banaras 1968 ; Kolhapur 73)**

10. If addition and multiplication modulo 10 is defined on the set of integers $R=\{0, 2, 4, 6, 8\}$, prove that the resulting system is a ring with unity. Is it an integral domain ? **Ans.** Yes.

11. Let A be the set of all real valued functions on $(-\infty, \infty)$. Define $(f+g)\,(x)=f\,(x)+g\,(x)$ and $(f\times g)\,(x)=f\,(g\,(x))$ for every x in $(-\infty, \infty)$. Is A a ring with respect to these two operations ?

(Kerala 1969)

Ans. No ; the distributive laws do not hold.

12. Show that a ring R which has a unique element e with the property that $ea=a\ \forall\ a \in R$, has unity. **(Meerut 1979)**

13. If a ring R has a left identity as well as right identity, then the two are equal. **(Lucknow 1968)**

14. Prove that the only idempotent elements of an integral domain with unity are 0 and 1. **(Nagarjuna 1978)**

15. An element a of a ring R is said to be **nilpotent** if $a^n=0$ for some positive integer n. Prove that $a=0$ is the only nilpotent element of an integral domain.

16. (i) Prove that a ring R is commutative if and only if $(a+b)^2=a^2+2ab+b^2\ \forall\ a, b \in R$. **(Agra 1980)**

(ii) If R is a commutative ring, prove by induction that $(a+b)^n=a^n+{}^nc_1a^{n-1}b+{}^nc_2a^{n-2}b^2+\ldots+b^n$ for every positive integer n ; here a and b are elements of R. **(Allahabad 1965)**

17. Prove that every field is an integral domain but every

integral domain is not a field. Give an example of an integral domain which is also a field. (Meerut 1977)

18. Prove that the set of matrices

$$\begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$$

where a, b are complex numbers and $\bar{a}, \bar{b}$ denote the complex conjugates of a and b, is a skew field for matrix addition and matrix multiplication.

19. Give examples to illustrate the difference between a field, a skew field and an integral domain. (Meerut 1978, 79, 82)

20. Define a ring and an integral domain. Give an example of a ring which is not an integral domain. (Marathwada 1970)

21. Give an example each of :

(i) a non-commutative ring. (Guru Nanak 1975)

(ii) Ring without zero divisors.

(iii) divison-ring. (Lucknow 1970 ; Marathwada 75)

(iv) A ring which is not an integral domain. (Meerut 1984)

22. Explain with an example the difference between a field and a skew field. (Meerut 1976)

23. If in a ring with unity any element a has the multiplicative inverse, then a cannot be a divisor of zero. (Lucknow 1969)

24. Prove that in a field :

(i) $\frac{a}{b}=\frac{c}{d} \Leftrightarrow ad=bc$. (ii) $\frac{a}{b}-\frac{c}{d}=\frac{ad-bc}{bd}$.

(iii) $(-a)^{-1}=-(a^{-1})$. (iv) $\frac{(-a)}{(-b)}=\frac{a}{b}$. (Allahabad 1970)

25. Show that $Z[\sqrt{-5}]$, the set of complex numbers $a+b\sqrt{(-5)}$ where a, b are integers, is an integral domain. (Meerut 1974)

26. Prove that the ring $I/(p)$, of integers modulo p is a field if and only if p is prime. (I.A.S. 1974)

27. Define commutative rings, integral domains and field. For the following, tell which of the above these are :

(i) The ring of integers modulo 4.

(ii) The ring of all 2×2 matrices over rationals.

(iii) The ring of all complex numbers.

(iv) The ring of rational numbers of the type a/p^n, where p

is a fixed prime, a any integer and n any integer $\geqslant 0$.

(Meerut 1976)

Ans. (i) commutative ring ; (ii) none ; (iii) field ;
(iv) integral domain.

28. Show that the set of all 2×2 matrices over a field F forms a non-commutative ring under matrix addition and multiplication. Show also that this ring contains divisors of zero.

(I.A.S. 1971 ; Meerut 77)

29. Let R be a ring with unit element. Let $\overline{R}$ denote an algebraic system consisting of the same elements as R and the two compositions of $\oplus$ and $*$, called, respectively, addition and multiplication, defined in it by the rule :

$a \oplus b = a + b + 1$, and $a * b = ab + a + b$ where $a, b \in R$

and where the addition and multiplication on the right hand side of these relations are those of R. Prove that $\overline{R}$ is a ring. Find the zero element and the unit element of $\overline{R}$.

(I.A.S. 1974 ; Raj. 78 ; Madras 78)

30. Show that the set of all 2×2 non-singular matrices over rationals is not a ring under matrix addition and multiplication.

(Meerut 1973)

31. Prove the cancellation law from the other postulates for a field. **(Kolhapur 1973)**

32. Enumerate the properties of a set to define an integral domain. **(I.A.S. 1971)**

33. If R is any ring, express $(a+b)^3$ as a sum of products of a and b, $a, b \in R$. **(Poona 1973)**

34. For what values of n, is $\mathbf{I}_n$,

(i) integral domain (ii) field ? **(Poona 1973)**

35. Give an example of a field F, which satisfies $Q \subset F \subset R$, where Q and R are the fields of rational and real numbers respectively.

36. Show that a finite ring with unity and no divisors of zero is a division ring. **(Meerut 1980)**

§ 7. Isomorphism of Rings. *A ring R is said to be isomorphic to another ring R′ if there exists a one-one mapping f of R onto R′ such that*

$$f(a+b) = f(a) + f(b),\ f(ab) = f(a)\, f(b)\ \forall\ a, b \in R.$$

Also such a mapping f is said to be an isomorphism of R onto R′.

A mapping $f : R \to R'$ is an isomorphism of R into R' if f is one-one and if $f(a+b)=f(a)+f(b)$ and $f(ab)=f(a)\,f(b)\ \forall\ a, b \in R$. For f to be an isomorphism of R onto R', it must be one-one onto and it must preserve both addition and multiplication compositions in R and R'. There may be more than one isomorphisms of R onto R' if R is isomorphic to R'.

If a ring R is isomorphic to another ring R', we shall write in symbols $R \cong R'$. Also R' is said to be an **isomorphic image of** R.

Note. In the above definition of isomorphism of rings, we have denoted the compositions in the two rings by the same symbols. The elements tell us that what is the composition represented by any symbol. For example, $a, b \in R$. When we write $a+b$, ab then the respective compositions are addition and multiplication of R. Again $f(a), f(b) \in R'$. When we write $f(a)+f(b)$, $f(a)\,f(b)$ then the respective compositions are addition and multiplication of R'.

Relation of isomorphism in the set of all rings.

As we have proved in groups, we can prove that the *relation of isomorphism in the set of all rings is an equivalence relation.* Therefore it will partition the set of all rings into disjoint equivalence classes such that any two rings of the same class are isomorphic to each other while rings of different classes are not isomorphic. Any two rings belonging to the same equivalence class are said to be **abstractly identical.**

Example. *Let R be the ring of integers under ordinary addition and multiplication. Let R′ be the set of all even integers. Let us define multiplication in R′ to be denoted by '∗' by the relation*

$$a * b = \frac{ab}{2},$$

where ab is the ordinary multiplication of two integers a and b.

(*i*) *Prove that* $(R', +, *)$ *is a commutative ring where + stands for ordinary addition of integers.*

(*ii*) *Prove that R is isomorphic to R′.*

(*iii*) *What acts as the unit element of R′ ?*

Solution. Obviously R' is an abelian group with respect to addition.

If a and b are both even integers then $\frac{ab}{2}$ is also an even integer. Therefore $a * b = \left(\frac{ab}{2}\right) \in R'\ \forall\ a, b \in R'$.

Thus R' is closed with respect to $*$.

Also if a, b, c are any elements of R', then

$$a * (b * c) = a * \left(\frac{bc}{2}\right) = \frac{a\,(bc/2)}{2} = \frac{(ab/2)\,c}{2} = \left(\frac{ab}{2}\right) * c = (a * b) * c.$$

$\therefore$ $*$ is associative.

Further $(a * b) = \frac{ab}{2} = \frac{ba}{2} = (b * a)$.

$\therefore$ $*$ is commutative.

Again $a * (b+c) = \frac{a\,(b+c)}{2} = \frac{ab}{2} + \frac{ac}{2} = (a*b) + (a*c)$.

Similarly $(b+c) * a = (b*a) + (c*a)$.

$\therefore$ $*$ is distributive with respect to $+$.

$\therefore$ $(R', +, *)$ is a commutative ring.

(ii) We shall now show that R is isomorphic to R'. Consider the mapping f defined by

$$f : R \to R' \text{ such that } f(x) = 2x \ \forall\ x \in R.$$

The mapping f is obviously one-one onto.

Also $\forall\ x_1, x_2 \in R$ we have

$$f(x_1 + x_2) = 2\,(x_1 + x_2) = 2x_1 + 2x_2 = f(x_1) + f(x_2)$$

and $f(x_1\, x_2) = 2\,(x_1\, x_2) = \frac{(2x_1)\,(2x_2)}{2} = (2x_1) * (2x_2) = f(x_1) * f(x_2)$.

$\therefore$ the mapping f is an isomorphism of R onto R'.

(iii) Here 1 is the unit element of R. We have $f(1) = 2$. Since f is an isomorphism of R onto R', therefore 2 must be the unit element of R'. We have for all $a \in R'$, $2 * a = \frac{2a}{2} = a = a * 2$.

$\therefore$ 2 is the unit element of R'.

§ 8. Properties of isomorphism of rings.

Theorem. *If f is an isomorphism of a ring R onto a ring R', then*

(*i*) *the image of the zero of R is the zero of R'.* (Nagpur 75)

(*ii*) *the image of the negative of an element of R is the negative of the image of that element i.e., $f(-a) = -f(a)\ \forall\ a \in R$.* (Nagpur 75)

(*iii*) *If R is commutative ring, then R' is also a commutative ring.*

(*iv*) *If R is without zero divisors, then R' is also without zero divisors.*

(*v*) *If R is with unit element, then R' is also with unit element.*

(*vi*) *If R is a field, then R' is also a field.* **(Kanpur 1980)**

(*vii*) *If R is a skew field, then R' is also a skew field.*

Proof. (i) Let $a \in R$. Then $f(a) \in R'$. Let $0'$ denote the zero element of R'. To prove that $f(0)=0'$.

We have $f(a)+0'=f(a)=f(a+0)=f(a)+f(0)$. By cancellation law for addition in R', we get from $f(a)+0'=f(a)+f(0)$, the result that $0'=f(0)$.

(ii) We have $f(a)+f(-a)=f[a+(-a)]=f(0)=0'$.

$\therefore$ $f(-a)$ is the additive inverse of $f(a)$ in R'. Thus

$$f(-a)=-f(a).$$

(iii) Let $f(a)$ and $f(b)$ be any two elements of R'. Then

$$a, b \in R.$$

We have $f(a) f(b)=f(ab)=f(ba)$

$$[\because R \text{ is commutative} \Rightarrow ab=ba]$$

$$=f(b) f(a).$$

$\therefore$ R' is also commutative.

(iv) We have $f(0)=0'$. Also f is one-one. Therefore 0 is the only element of R whose f-image is $0'$.

Let $f(a), f(b)$ be two non-zero elements of R'. Then $f(a) \neq 0'$, $f(b) \neq 0' \Rightarrow a \neq 0, b \neq 0$. Since R is without zero divisors, therefore

$a \neq 0, b \neq 0 \Rightarrow ab \neq 0 \Rightarrow f(ab) \neq f(0)$

$\Rightarrow f(a) f(b) \neq 0' \Rightarrow R'$ is without zero divisors.

(v) Let 1 be the unit element of R. Then $f(1) \in R'$. If $f(a)$ is any element of R', we have

$f(1) f(a)=f(1a)=f(a)$ and $f(a) f(1)=f(a1)=f(a)$.

$\therefore$ $f(1)$ is the unit element of R'.

(vi) If R is a field, then R is commutative, with unity and each non-zero element of R will possess multiplicative inverse. Now as proved in (iii) and (v), R' will be commutative and will also have the unit element *i.e.*, $f(1)$.

Let $f(a)$ be any non-zero element of R'. Then

$$f(a) \neq 0' \Rightarrow a \neq 0 \Rightarrow a^{-1} \text{ exists.}$$

Now $f(a^{-1}) \in R'$ and we have

$f(a^{-1}) f(a)=f(a^{-1} a)=f(1)$ and $f(a) f(a^{-1})=f(aa^{-1})=f(1)$.

$\therefore$ $f(a^{-1})$ is the multiplicative inverse of $f(a)$.

Hence R' is a field.

(vii) As shown in (v) R' will be with unit element *i.e.*, $f(1)$ and as shown in (vi) each non-zero element of R' will be inversible. Therefore R' is a skew-field.

Transference of ring structure.

Theorem. *If f is a one-one mapping of a ring R onto a set R' with two compositions denoted additively and multiplicatively such that $f(a+b)=f(a)+f(b)$, $f(ab)=f(a)\,f(b)$ $\forall$ $a, b \in R$, then the set R' is a ring for the two compositions.* **(Banaras 1967)**

The theorem can be easily proved as we have proved the corresponding theorem on groups.

§ 9. Subrings.

Definition. *Let R be a ring. A non-empty subset S of the set R is said to be a subring of R if S is closed with respect to the operations of addition and multiplication in R and S itself is a ring for these operations.* **(Indore 1970; Meerut 74)**

If S is a subring of a ring R, it is obvious that S is a subgroup of the additive group of R.

If R is any ring, then $\{0\}$ and R itself are always subrings of R. These are known as **improper subrings** of R. Other subrings if any, of R are called **proper subrings** of R.

Conditions for a subring. Theorem. *The necessary and sufficient conditions for a non-empty subset S of a ring R to be a subring of R are*

(*i*) $a \in S, b \in S \Rightarrow a-b \in S$ (*ii*) $a \in S, b \in S \Rightarrow ab \in S$.

(Indore 1970; Madurai 78; Meerut 80, 81; Garhwal 76; Rajasthan 76)

Proof. The conditions are necessary. Suppose $(S, +, \cdot)$ is a subring of $(R, +, .)$.

Since S is a group with respect to addition, therefore $b \in S \Rightarrow -b \in S$.

Now S is closed with respect to addition.

$$\therefore \quad a \in S, b \in S \Rightarrow a \in S, -b \in S$$
$$\Rightarrow a+(-b) \in S \Rightarrow a-b \in S.$$

Also S is closed with respect to multiplication.

$$\therefore \quad a \in S, b \in S \Rightarrow ab \in S.$$

Hence the conditions are necessary.

The conditions are sufficient. Suppose S is a non-empty subset of R and the conditions (i) and (ii) are satisfied. From (i), we have

$$a \in S, a \in S \Rightarrow a-a \in S$$
$$\Rightarrow 0 \in S$$ *i.e.*, the zero element $\in S$.

Now since $0 \in S$, therefore from (i), we have

$0 \in S, a \in S \Rightarrow 0-a \in S \Rightarrow -a \in S$ *i.e.*, each element of S possesses additive inverse.

Now if a, b are any elements of S, then $-b \in S$.

From (i), we have

$a \in S, -b \in S \Rightarrow a-(-b) \in S \Rightarrow a+b \in S$.

$\therefore$ S is closed with respect to addition.

Now S is a subset of R. Therefore associativity and commutativity of addition must hold in S since they hold in R.

$\therefore$ $(S, +)$ is an abelian group.

From (ii) S is closed with respect to multiplication.

Associativity of multiplication and distributivity of multiplication over addition must hold in S since they hold in R.

Hence S is a subring of R.

Cor. *The necessary and sufficient conditions for a non-empty subset S of a ring R to be a subring of R are*

(*i*) $S+(-S)=S$ (*ii*) $SS \subseteq S$. **(Meerut 1970)**

Intersection of subrings.

Theorem 1. *The intersection of two subrings is a subring.*

(Indore 1970, Meerut 84, Kanpur 80, Garhwal 76; Rajasthan 77)

Proof. Let S_1 and S_2 be two subrings of a ring R. Then $S_1 \cap S_2$ is not empty since at least $0 \in S_1 \cap S_2$.

Now in order to prove that $S_1 \cap S_2$ is a subring, it is sufficient to prove that

(*i*) $a \in S_1 \cap S_2, b \in S_1 \cap S_2 \Rightarrow a-b \in S_1 \cap S_2$

and (ii) $a \in S_1 \cap S_2, b \in S_1 \cap S_2 \Rightarrow ab \in S_1 \cap S_2$.

We have $a \in S_1 \cap S_2 \Rightarrow a \in S_1, a \in S_2$

$b \in S_1 \cap S_2 \Rightarrow b \in S_1, b \in S_2$.

Now S_1 and S_2 are both subrings.

$\therefore$ $a \in S_1, b \in S_1 \Rightarrow a-b \in S_1$ and $ab \in S_1$

and $a \in S_2, b \in S_2 \Rightarrow a-b \in S_2$ and $ab \in S_2$.

Now $a-b \in S_1, a-b \in S_2 \Rightarrow a-b \in S_1 \cap S_2$

and $ab \in S_1, ab \in S_2 \Rightarrow ab \in S_1 \cap S_2$.

Thus $a \in S_1 \cap S_2, b \in S_1 \cap S_2$

$\Rightarrow a-b \in S_1 \cap S_2$ and $ab \in S_1 \cap S_2$.

$\therefore$ $S_1 \cap S_2$ is a subring of R.

Theorem 2. *An arbitrary intersection of subrings is a subring.*

Proof. Let R be a ring and let $\{S_t : t \in T\}$ be any family of subrings of R. Here T is an index set and is such that $\forall\ t \in T$,

S_t is a subring of R. Let $S = \bigcap_{t \in T} S_t = \{x \in R : x \in S_t \ \forall \ t \in T\}$ be the intersection of this family of subrings of R. Then to prove that S is also a subring of R.

Obviously $S \neq \emptyset$, since at least the zero element 0 of R is in $S_t \ \forall \ t \in T$.

Now let a, b be any two elements of S. Then

$$a \in \bigcap_{t \in T} S_t \Rightarrow a \in S_t \ \forall \ t \in T$$

and

$$b \in \bigcap_{t \in T} S_t \Rightarrow b \in S_t \ \forall \ t \in T.$$

But $\forall \ t \in T$, S_t is a subring of R. Therefore

$$a, b \in S_t \Rightarrow a-b, ab \in S_t \ \forall \ t \in T.$$

Consequently $a-b, ab \in \bigcap_{t \in T} S_t$. Thus we have shown that

$$a, b \in \bigcap_{t \in T} S_t \Rightarrow a-b, ab \in \bigcap_{t \in T} S_t.$$

Therefore $\bigcap_{t \in T} S_t$ is a subring of R.

Smallest subring containing a given subset of a ring. Suppose R is a ring and M is any subset of R. Further suppose that S is a subring of R such that $M \subseteq S$ and if T is any subring of R containing M then $S \subseteq T$. Then S is called the subring of R generated by the subset M. In short if S is the smallest subring of R containing M, then S is called the subring generated by M. Also we write $S = (M)$.

Theorem. *The intersection of the family of subrings which contain a given subset M of a ring R is the smallest subring containing the subset M.*

Proof. The family of subrings which contain M is not empty since at least R is a subring of R which contains M. Further the intersection of all subrings of R which contain M is also a subring of R and M is contained in this intersection. Also this intersection will be contained in any subring of R which contains M. Therefore this intersection is the smallest subring of R containing M *i.e.*, it is the subring of R generated by M.

Some Examples of subrings.

Example 1. *The set of integers is a subring of the ring of rational numbers.*

If $a, b \in I$, then $a-b \in I$ and $ab \in I$.

∴ I is a subring of the ring of rational numbers.

Example 2. *The set of all $n \times n$ matrices over the field of rational numbers is a subring of the ring of all $n \times n$ matrices over the field of real numbers.*

Example 3. *Let R be the ring of all 2×2 matrices over the field of real numbers. Let M be a subset of R and let the elements of M be matrices of the type*

$$\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix}$$

i.e., matrices in which the elements of second column are all zero. Then M is a subring of R.

Let $A=\begin{bmatrix} a_1 & 0 \\ b_1 & 0 \end{bmatrix}$, $B=\begin{bmatrix} a_2 & 0 \\ b_2 & 0 \end{bmatrix}$, be any two elements of M.

Then $A-B=\begin{bmatrix} a_1-a_2 & 0 \\ b_1-b_2 & 0 \end{bmatrix}$. Also $AB=\begin{bmatrix} a_1 & 0 \\ b_1 & 0 \end{bmatrix}\begin{bmatrix} a_2 & 0 \\ b_2 & 0 \end{bmatrix}=\begin{bmatrix} a_1a_2 & 0 \\ b_1a_2 & 0 \end{bmatrix}$.

Now $A-B$ and AB are both members of M since the second column of $A-B$ and also of AB consists of zeros only.

∴ M is a subring of R.

Example 4. *Show that the set of matrices $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$ is a subring of the ring of 2×2 matrices with integral elements.*

(Meerut 1979 ; Madras 83)

Solution. Let R be the ring of 2×2 matrices with integral elements and let M be the subset of R and let the elements of M be matrices of the type $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$.

Let $A=\begin{bmatrix} a_1 & b_1 \\ 0 & c_1 \end{bmatrix}$, $B=\begin{bmatrix} a_2 & b_2 \\ 0 & c_2 \end{bmatrix}$ be any two elements of M.

Then $A-B=\begin{bmatrix} a_1-a_2 & b_1-b_2 \\ 0 & c_1-c_2 \end{bmatrix}$ which is obviously an element of M.

Also $AB=\begin{bmatrix} a_1 & b_1 \\ 0 & c_1 \end{bmatrix}\begin{bmatrix} a_2 & b_2 \\ 0 & c_2 \end{bmatrix}=\begin{bmatrix} a_1a_2 & a_1b_2+b_1c_2 \\ 0 & c_1c_2 \end{bmatrix}$ which is obviously an element of M.

∴ M is a subring of R.

Example 5. *Let R be the ring of integers. Let m be any fixed integer and let S be any subset of R such that*

$S=\{\ldots, -3m, -2m, -m, 0, m, 2m, 3m, \ldots\}$. Then S is a subring of R.

Let $a=rm$ and $b=sm$ be any two elements of S. Then r and s are some integers. We have

$a-b=rm-sm=(r-s)\,m$ and $ab=(rm)\,(sm)=(rsm)\,m$.

Since $r-s$ is some integer and (rsm) is also some integer, therefore both $a-b$ and ab are elements of S. Hence S is a subring of R.

§ 10. Subfields.

Definition. *Let F be a field. A non-empty subset K of the set F is said to be a subfield of F if K is closed with respect to the operations of addition and multiplication in F and K itself is a field for these operations.* **(Kanpur 1969)**

Conditions for a subfield. Theorem. *The necessary and sufficient conditions for a non-empty subset K of a field F to be a subfield of F are*

(*i*) $a \in K, b \in K \Rightarrow a-b \in K$.

(*ii*) $a \in K, 0 \neq b \in K \Rightarrow ab^{-1} \in K$. **(Bombay 1970)**

Proof. The conditions are necessary. Suppose K is a subfield of the field F. Now K is a group with respect to addition. Therefore $b \in K \Rightarrow -b \in K$. Also K is closed with respect to addition.

$$\therefore \quad a \in K, b \in K \Rightarrow a+(-b) \in K \Rightarrow a-b \in K.$$

Now each non-zero element of K possesses multiplicative inverse. Therefore $0 \neq b \in K \Rightarrow b^{-1} \in K$.

But K is closed with respect to multiplication.

$$\therefore \quad a \in K, 0 \neq b \in K \Rightarrow ab^{-1} \in K.$$

Hence the conditions are necessary.

The conditions are sufficient. Suppose K is a non-empty subset of F and the conditions (i) and (ii) are satisfied.

As we have proved in subrings, we can prove that with the help of condition (i), $(K, +)$ is an abelian group. [Give the same proof here].

Now let a be any non-zero element of K. Then from (ii) we have $a \in K, 0 \neq a \in K \Rightarrow aa^{-1} \in K \Rightarrow 1 \in K$.

Now $1 \in K$, therefore again from (ii), we have

$$1 \in K, 0 \neq a \in K \Rightarrow 1a^{-1} \in K \Rightarrow a^{-1} \in K.$$

$\therefore$ Each non-zero element of K possesses multiplicative inverse.

Now let $a \in K$ and $0 \neq b \in K$. Then $b^{-1} \in K$.

From (ii), we have

$$a \in K, 0 \neq b^{-1} \in K \Rightarrow a\,(b^{-1})^{-1} \in K \Rightarrow ab \in K.$$

Also if $b=0$, then $ab=0$ and $0 \in K$.

$$\therefore \quad ab \in K \ \forall \ a, b \in K.$$

Associativity of multiplication and distributivity of multiplication over addition must hold in K since they hold in F.

Hence K is a subfield of F.

Example. The field of real numbers is a subfield of the field of complex numbers. The field of rational numbers is a subfield of the field of real numbers.

§ 11. Characteristic of a ring.

Definition. *Let R be a ring with zero element 0 and suppose there exists a positive integer n such that $na=a+a+\ldots$upto n terms$=0$ for every $a \in R$. The smallest such positive integer n is called the characteristic of the ring R. If there exists no such positive integer, then R is said to be of characteristic zero or infinite.*

(Nagarjuna 1978; Rajasthan 76)

If any element of a ring R is of order zero when regarded as an element of the additive group $(R, +)$, then R will be of zero characteristic.

The ring of integers is of characteristic zero. The ring of rational numbers is also of characteristic zero.

If $I_6=\{0, 1, 2, 3, 4, 5\}$, then the ring $(I_6, +_6, \times_6)$ *i.e.*, the ring of integers modulo 6 has characteristic 6 since $6x=0$ for every x in the ring. Obviously no integer smaller than 6 satisfies this property. For instance, 5 cannot be the characteristic, since $5(2)=4$ in I_6 and $4 \neq 0$.

Theorem 1. *The characteristic of a ring with unity is 0 or $n>0$ according as the unity element 1 regarded as a member of the additive group of the ring has the order zero or n.*

Proof. Let R be a ring with unity element 1. If 1 has order zero, then the characteristic of the ring is zero.

Suppose 1 is of finite order n so that

$$1+1+1+\ldots\text{upto } n \text{ terms}=0 \text{ i.e., } n1=0.$$

Let a be any element of R. Then, we have

$$\begin{aligned} na &= a+a+a+\ldots\text{upto } n \text{ terms} \\ &= 1a+1a+1a+\ldots\text{upto } n \text{ terms} \\ &= (1+1+1+\ldots\text{upto } n \text{ terms})\,a && \text{[by dist. law]} \\ &= (n1)\,a=0a=0. \end{aligned}$$

$\therefore$ order of a is $\leqslant n$.

Hence the characteristic of the ring is n.

Theorem 2. *The characteristic of an integral domain is* 0 *or* $n > 0$ *according as the order of any non-zero element regarded as a member of the additive group of the integral domain is either* 0 *or* n.
(Rajasthan 1976)

Proof. Let D be an integral domain.

If a non-zero element of D is of order zero, then the characteristic of D is zero.

Let the order of the non-zero element a be finite and equal to n. Then $na=0$.

Suppose b is any other non-zero element of D.

We have $na=0$

$\Rightarrow (na)\, b=0$

$\Rightarrow (a+a+a+\ldots\text{upto } n \text{ terms})\, b=0$

$\Rightarrow (ab+ab+ab+\ldots\text{upto } n \text{ terms})=0$

$\Rightarrow a\,(b+b+b+\ldots\text{upto } n \text{ terms})=0$

$\Rightarrow a\,(nb)=0.$

But D is without zero divisors. Therefore $a \neq 0$ and $a\,(nb)=0 \Rightarrow nb=0$.

But the order of a is $n \Rightarrow n$ is the least positive integer such that $na=0$. Also we have $n0=0$. Thus n is the least positive integer such that $nx=0 \;\forall\; x \in D$. Hence D is of characteristic n.

Theorem 3. *Each non-zero element of an integral domain D, regarded as a member of the additive group of D, is of the same order.*
(Andhra 1975, 77)

Proof. Let D be an integral domain. Suppose a is a non-zero element of D and $o\,(a)$ is finite and say, equal to n.

Suppose b is any other non-zero element of D and $o\,(b)=m$.

We have $o\,(a)=n \Rightarrow na=0$

$\Rightarrow nb=0$ [See Theorem 2]

$\Rightarrow o\,(b) \leqslant n \Rightarrow m \leqslant n.$

Similarly $o(b)=m \Rightarrow mb=0 \Rightarrow a\,(mb)=0$

$\Rightarrow a\,(b+b+\ldots\text{upto } m \text{ times})=0$

$\Rightarrow (ab+ab+ab+\ldots\text{upto } m \text{ times})=0$

$\Rightarrow (a+a+a\ldots\text{upto } m \text{ times})\, b=0$

$\Rightarrow (ma)\, b=0$

$\Rightarrow ma=0$ [$\because$ $b \neq 0$ and D is without zero divisors]

$\Rightarrow o\,(a) \leqslant m \Rightarrow n \leqslant m.$

Now $m \leqslant n,\ n \leqslant m \Rightarrow m=n$. Hence $o(a)=o(b)$.

Also if $o\,(a)$ is zero, then $o\,(b)$ cannot be finite. Because $o\,(b)=m \Rightarrow ma=0$ *i.e.*, the order of a is finite. Hence $o\,(b)$ must also be zero. Hence the theorem.

Theorem 4. *The characteristic of an integral domain is either* 0 *or a prime number.* **(I.A.S. 1971; Nagarjuna 78, 79, 80; Kerala 70; Madras 77; Andhra 75, 77; Kanpur 80)**

Proof. Suppose D is an integral domain. Let $0 \neq a \in D$. If $o\,(a)$ is zero, then the characteristic of D is 0. If $o\,(a)$ is finite, let $o\,(a)=p$. Then the characteristic of D will be p. We are to prove that p must be prime.

Suppose p is not prime. Let $p=p_1p_2$ where $p_1 \neq 1$, $p_2 \neq 1$ and $p_1 < p$ also $p_2 < p$.

Since D is an integral domain, therefore the product of two non-zero elements of D cannot be equal to 0.

$$\therefore \quad aa \neq 0 \ \ i.e.,\ a^2 \neq 0.$$

Now in an integral domain two non-zero elements are of the same order.

$$\begin{aligned}
\therefore \quad o\,(a)=p &\Rightarrow o\,(a^2)=p \Rightarrow pa^2=0 \\
&\Rightarrow (p_1p_2)a^2=0 && [\because \ p=p_1p_2] \\
&\Rightarrow (a^2+a^2+a^2+\ldots \text{upto } p_1p_2 \text{ terms})=0 \\
&\Rightarrow (p_1a)\,(p_2a)=0 \\
&\Rightarrow \text{either } p_1a=0 \text{ or } p_2a=0 \\
& \qquad\qquad [\because \ D \text{ is without zero divisors}]
\end{aligned}$$

But $p_1 < p$ and $p_2 < p$. Also p is the least positive integer such that $pa=0$. Hence p must be prime.

Characteristic of a field. *Every field is an integral domain. Therefore the characteristic of a field F is* 0 *or* $n > 0$ *according as any non-zero element (in particular the unit element* 1*) of F is of order* 0 *or* n.

Thus in order to find the characteristic of a field F, we should find the order of the unit element 1 of F when regarded as a member of the additive group of F. If the order of 1 is zero, then F is of characteristic 0. If the order of 1 is finite, say, n then the characteristic of F is n.

The characteristic of the field of real numbers is 0.

The characteristic of the finite field $(I_7, +_7, \times_7)$ is 7 where $I_7=\{0, 1, 2, 3, 4, 5, 6\}$.

§ 12 Ordered Integral Domains.

Definition. *An integral domain* $(D, +, \cdot)$ *is said to be ordered if* D *contains a subset* D_+ *such that*

(*i*) D_+ *is closed with respect to addition and multiplication as defined on* D.

(*ii*) $\forall\ a \in D$, *one and only one of* $a=0$, $a \in D_+$, $-a \in D_+$, *holds* (*principle of Trichotomy*). **(Allahabad 1965, 70; Bombay 70)**

The elements of D_+ are called the *positive* elements of D; all other non-zero elements of D are called *negative* elements of D.

Ordered field *A field* $(F, +, \cdot)$ *is said to be ordered if it is ordered as an integral domain.* **(Kolhapur 1973)**

The integral domain $(\mathbf{I}, +, \cdot)$ of all integers is ordered.

The set $\mathbf{I}_+$ of all positive integers is the set of the positive elements of this integral domain. We know that the sum and product of two positive integers is again a positive integer *i.e.*, $\mathbf{I}_+$ is closed with respect to addition and multiplication. If $a \in \mathbf{I}$, then either a is zero or positive or negative *i.e.*, either $a=0$ or $a \in \mathbf{I}_+$ or $a \in \mathbf{I}_+$.

The field of rational numbers is an ordered field. The field of real numbers is also an ordered field. But the field of complex numbers is not an ordered field.

Theorem 1. *Let* D *be an integral domain with unity element* 1. *If* D *is an ordered integral domain show that* 1 *is a positive element of* D. **(Allahabad 1970)**

Proof. Let D be an ordered integral domain with unity element 1. Let D_+ denote the set of positive elements of D.

Suppose $1 \notin D_+$.

Now $1 \neq 0$. Since $1 \notin D_+$ therefore by the definition of an ordered integral domain,

$$-1 \in D_+$$

$$\Rightarrow (-1)(-1) \in D_+ \qquad [\because D_+ \text{ is closed with respect to multiplication}]$$

$$\Rightarrow 1 \in D_+ \text{ which is a contradiction.}$$

Hence $1 \in D_+$ *i.e.*, 1 is a positive element of D.

Theorem 2. *The field* $(\mathbf{C}, +, \cdot)$ *of complex numbers is not ordered.* **(Banaras 1969; Meerut 76, 78; Allahabad 70)**

Proof. Suppose $\mathbf{C}$ is an ordered field and $\mathbf{C}_+$ is the set of positive elements of this field. The additive identity *i.e.*, the zero element is $\quad 0+i0$.

Now $\quad i \neq 0$.

By the principle of trichotomy either $i \in C_+$ or $-i \in C_+$.

Now C_+ is closed with respect to multiplication

$\therefore\ i \in C_+ \Rightarrow i.i = -1 \in C_+$.

Again $i \in C_+$, $-1 \in C_+ \Rightarrow i(-1) = -i \in C_+$. Thus if $i \in C_+$, its additive inverse $-i$ also belongs to C_+. This contradicts the principle of trichotomy *i.e.*, the condition (ii) of the definition of an ordered integral domain.

Similarly if we assume that $-i \in C_+$, we can show that its additive inverse i also belongs to C_+. This again contradicts the principle of trichotomy.

Hence the field of complex numbers is not an ordered field.

Theorem 3. *The field* $(I_p, +_p, \times_p)$ *where* p *is a prime and* $I_p = \{0, 1, \ldots, p-1\}$ *is not ordered.*

Proof. Suppose I_p is an ordered field and P is the set of positive elements of this field. The zero element of this field is 0.

Now $1 \neq 0$. The additive inverse of 1 is $p-1$. According to the definition of ordered field either $1 \in P$ or its additive inverse $p-1 \in P$.

But P is closed with respect to $+_p$.

Therefore $1 \in P \Rightarrow 1 +_p 1 +_p 1 +_p \ldots$ upto $p-1$ times $\in P$

$\Rightarrow p-1 \in P$.

This contradicts the principle of trichotomy. Similarly assuming that $p-1 \in P$ we can show that its additive inverse 1 also belongs to P. This again contradicts the principle of trichotomy. Hence the given field is not an ordered field.

Order relations in an ordered integral domain.

Definition. *Let D be an ordered integral domain and D_+ be the set of positive elements of D. Then we define 'less than' ($<$) 'greater than' ($>$) relations in D as follows :*

For all $a, b \in D$, we have

(i) $a > b$ *when* $a - b \in D_+$.

(ii) $a < b$ *when* $b - a \in D_+$.

Obviously $a > b$ iff $b < a$.

Theorem. *The order relation in an ordered integral domain is transitive i.e.,* $a > b,\ b > c \Rightarrow a > c$. **(Kolhapur 73)**

Proof. Let D be an ordered integral domain and let D_+ be the set of positive elements of D.

We have $a > b \Rightarrow a - b \in D_+$ [by def. of $>$]

and $b > c \Rightarrow b - c \in D_+$.

Now D_+ is closed with respect to addition.

$$\therefore \quad a-b \in D_+, b-c \in D_+ \Rightarrow (a-b)+(b-c) \in D_+$$
$$\Rightarrow a-c \in D_+ \Rightarrow a > c.$$

Exercises

1. If R is a ring, show that $Z(R)=\{x \in R : xy=yx, \forall y \in R\}$ is a subring of R. Further show that $Z(R)$ is a field if R is a division ring. **(Delhi 1969, 70)**

2. Prove or disprove that any subring of a non-commutative ring is non-commutative. **(Meerut 1976)**

3. Let x, y be commutative elements of a ring R of characteristic two. Show that $(x+y)^2=x^2+y^2=(x-y)^2$.

4. Let R be a non-zero ring such that for all $a \in R$, $a^2=a$. Prove that R is a commutative ring of characteristic 2. **(Madras 1974; Nagarjuna 1978, 79, 80)**

5. Show that every finite integral domain is of finite characteristic.

6. Show that in an integral domain all non-zero elements generate additive cyclic groups of the same order which is equal to the characteristic of the integral domain. **(Allahabad 1978)**

7. Give without proof, an example of an integral domain which contains only five elements. Is this an ordered integral domain ? Give reason. **(Allahabad 1970)**

8. Define the characteristic of a ring and prove that if R is a finite ring then the characteristic of R is finite and $\neq 0$. **(Kerala 1970)**

9. Define an ordered field and illustrate the concept with the help of an example. **(Meerut 1974)**

10. Show that the set of Matrices $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, a, b real, forms a field isomorphic to the field of complex numbers.

11. Give an example of a ring with unity 1 which has a subring with unity $1' \neq 1$. **(Meerut 1981)**

§ 13. Imbedding of a ring into another ring.

Definition. *A ring R is said to be imbedded in a ring R' if there is a subring S' of R' such that R is isomorphic to S'.*

Obviously a ring R can be imbedded in a ring R' if there exists a mapping f of R into R' such that f is one-to-one and

$$f(a+b)=f(a)+f(b), f(ab)=f(a) f(b) \ \forall \ a, b \in R.$$

For then $f(R)$ is a subring of R' and f is an isomorphism of R onto $f(R)$ making R isomorphic to $f(R)$.

Theorem. *Any ring R without a unity element can be imbedded in a ring with unity.* **(Sagar 1966; Meerut 78, 79; Jabalpur 70)**

Proof. Let R be any ring without unity. Let Z be the ring of integers. Let $R'=R\times Z=\{(a, m) : a \in R \text{ and } m \in Z\}$.

We shall show that when suitable binary operations have been defined in $R\times Z$, then it becomes a ring with a unity element containing a subring, isomorphic to R.

If (a, m) and (b, n) are any two elements of $R\times Z$, then we define addition in $R\times Z$ by the equation

$$(a, m)+(b, n)=(a+b, m+n) \qquad \text{.. (1)}$$

and multiplication in $R\times Z$ by the equation

$$(a, m)(b, n)=(ab+na+mb, mn). \qquad \text{...(2)}$$

In the right hand side of (1), $a+b$ is addition of two elements of R and $m+n$ is addition of two integers. In the right hand side of (2), ab is multiplication of two elements of R, mn is multiplication of two integers and na, mb are integral multiples of a and b which we have explained in § 3 page 256 of this chapter.

Since $a+b \in R$ and $m+n \in Z$, therefore $(a+b, m+n) \in R\times Z$. Thus $R\times Z$ is closed with respect to addition. Further $ab, na, mb \in R \Rightarrow ab+na+mb \in R$. Also $mn \in Z$. Therefore $(ab+na+mb, mn) \in R\times Z$ and $R\times Z$ is closed with respect to multiplication.

Now let (a, m), (b, n), (c, p) be any elements of $R\times Z$. Then we observe :

Associativity of addition. We have

$$[(a, m)+(b, n)]+(c, p)=(a+b, m+n)+(c, p)$$
$$=([a+b]+c, [m+n]+p)=(a+[b+c], m+[n+p])$$
$$=(a, m)+(b+c, n+p)=(a, m)+[(b, n)+(c, p)].$$

Commutativity of addition. We have

$$(a, m)+(b, n)=(a+b, m+n)$$
$$=(b+a, n+m) \qquad [\because \text{ addition is commutative in } R \text{ and also in } Z]$$
$$=(b, n)+(a, m).$$

Existence of additive identity. We have $(0, 0) \in R\times Z$. Here the first 0 is the zero element of R and the second 0 is the zero integer. Also $(0, 0)+(a, m)=(0+a, 0+m)=(a, m)$.

$\therefore$ (0, 0) is the additive identity.

Existence of additive inverse. If $(a, m) \in R\times Z$, then

$(-a, -m) \in R\times Z$ and we have

$$(-a, -m)+(a, m)=(-a+a, -m+m)=(0, 0).$$

$\therefore$ $(-a, -m)$ is the additive inverse of (a, m).

Associativity of multiplication. We have

$$[(a, m)(b, n)](c, p)=(ab+na+mb, mn)(c, p)$$

$=((ab+na+mb)\,c+p\,(ab+na+mb)+(mn)\,c,\ (mn)\,p)$
$=(abc+n(ac)+m\,(bc)+p\,(ab)+(pn)\,a+(pm)\,b+(mn)\,c,\ (mn)\,p).$

Also $(a, m)\,[(b, n)\,(c, p)]=(a, m)\,(bc+pb+nc,\ np)$
$=(a\,(bc+pb+nc)+(np)\,a+m\,(bc+pb+nc),\ m\,(np))$
$=(abc+a\,(pb)+a\,(nc)+(np)a+m\,(bc)+m(pb)+m(nc),\ (mn)\,p)$
$=(abc+p(ab)+n(ac)+(np)a+m(bc)+(mp)\,b+(mn)\,c,\ (mn)\,p).$

We see that $[(a, m)\,(b, n)]\,(c, p)=(a, m)\,[(b, n)\,(c, p)]$.

Distributive laws. We have

$(a, m)\,[(b, n)+(c, p)]=(a, m)\,(b+c,\ n+p)$
$=(a\,(b+c)+(n+p)\,a+m\,(b+c),\ m\,(n+p))$
$=(ab+ac+na+pa+mb+mc,\ mn+mp)$
$=(ab+na+mb,\ mn)+(ac+pa+mc,\ mp)$
$=(a, m)\,(b, n)+(a, m)\,(c, p).$

Similarly we can show that the other distributive law also holds good.

Thus $R \times Z$ is a ring with respect to the operations defined on it.

Existence of multiplicative identity. We have

$(0, 1) \in R \times Z$. If $(a, m) \in R \times Z$, then

$(0, 1)\,(a, m)=(0a+m0+1a,\ 1m)=(0+0+a,\ m)=(a, m).$

Also $(a, m)\,(0, 1)=(a0+1a+m0,\ m1)=(0+a+0,\ m)=(a, m).$

$\therefore$ $(0, 1)$ is the multiplicative identity. So $R \times Z$ is a ring with unity element $(0, 1)$.

Now consider the subset $S'=R \times \{0\}$ of $R \times Z$ which consists of all pairs of the form $(a, 0)$. We shall show that $R \times \{0\}$ is a subring of $R \times Z$. Let $(a, 0)$, $(b, 0)$ be any two elements of $R \times \{0\}$. Then $(a, 0)-(b, 0)=(a, 0)+(-b, -0)=(a-b, 0-0)=(a-b, 0)$
$\in R \times \{0\}$.

Also $(a, 0)\,(b, 0)=(ab+0a+0b,\ 00)=(ab+0+0,\ 0)$
$=(ab, 0) \in R \times \{0\}$.

$\therefore$ $R \times \{0\}$ is a subring of $R \times Z$.

Finally we shall show that R is isomorphic to $R \times \{0\}$. Let ϕ be a mapping from R to $R \times \{0\}$ defined as $\phi\,(a)=(a, 0)\ \forall\ a \in R$.

ϕ is one-one. We have

$\phi(a)=\phi(b) \Rightarrow (a, 0)=(b, 0) \Rightarrow a=b \Rightarrow \phi$ is one-one.

ϕ is onto. Let $(a, 0)$ be any element of $R \times \{0\}$. Then $a \in R$ and we have $\phi\,(a)=(a, 0)$. Therefore ϕ is onto.

ϕ preserves additions and multiplications.

If a, b be any two elements of R, then

$$\phi\,(a+b)=(a+b, 0)=(a, 0)+(b, 0)=\phi\,(a)+\phi(b).$$

Also $\phi(ab)=(ab, 0)=(a, 0)(b, 0)=\phi(a)\phi(b)$.

$\therefore$ ϕ preserves compositions.

Hence ϕ is an isomorphism of R onto $R\times\{0\}$.

This completes the proof of the theorem.

§ 14. The field of Quotients.

Definition *A ring R can be imbedded in a ring S if S contains a subset S′ such that R is isomorphic to S′.*

If D is a commutative ring without zero divisors, then we shall see that it can be imbedded in a field F *i.e.*, there exists a field F which contains a subset D' isomorphic to D. We shall construct a field F with the help of elements of D and this field F will contain a subset D' such that D is isomorphic to D'. This field F is called the "**field of quotients**" of D, or simply the "*quotient field*" of D.

On account of isomorphism of D onto D', we can say that D and D' are abstractly identical. Therefore if we identify D' with D, then we can say that the quotient field F of D is a field containing D. We shall also see that F is the smallest field containing D.

Motivation for the construction of the quotient field. We are all quite familiar with the ring **I** of integers. Also our familiar set **Q** of rational numbers is nothing but the set of quotients of the elements of **I**. Thus $\mathbf{Q}=\left\{\frac{p}{q} : p\in\mathbf{I},\ 0\neq q\in\mathbf{I}\right\}$. If we identify the rational numbers..., $\frac{-3}{1}, \frac{-2}{1}, \frac{-1}{1}, \frac{0}{1}, \frac{1}{1}, \frac{2}{1}, \frac{3}{1}$... with the integers ..., $-3, -2, -1, 0, 1, 2, 3$,..., then $\mathbf{I}\subseteq\mathbf{Q}$.

Also $(\mathbf{Q}, +, \cdot)$ is a field. It is the smallest field containing **I**.

Also if $\frac{a}{b}$ and $\frac{c}{d}\in\mathbf{Q}$, then we remember that

(i) $\frac{a}{b}=\frac{c}{d}$ iff $ad=bc$, (ii) $\frac{a}{b}+\frac{c}{d}=\frac{ad+bc}{bd}$, (iii) $\frac{a}{b}\,\frac{c}{d}=\frac{ac}{bd}$.

Taking motivation from these facts, we now proceed to construct the quotient field of an arbitrary integral domain. We have the following theorem :

Theorem 1. *A commutative ring without zero divisors can be imbedded in a field.* **(I.A.S. 1969; Rajasthan 74; Andhra 75, 77; Nagarjuna 78, 80; Allahabad 80; Utkal 70; Banaras 70)**

Or

Every integral domain can be imbedded in a field. *Or*

From the elements of an integral domain D, it is possible to construct a field F which contains a subset D′ isomorphic to D.

Proof. Let D be a commutative ring without zero divisors. Let D_0 be the set of all non-zero elements of D. Let $S = D \times D_0$ *i.e.*, let S be the set of all ordered pairs (a, b) where $a, b \in D$ and $b \neq 0$. Let us define a relation $\sim$ in S. We shall say that

$$(a, b) \sim (c, d) \text{ if and only if } ad = bc.$$

We claim that this relation is an equivalence relation in S.

Reflexivity. Since D is a commutative ring, therefore $ab = ba$ $\forall\ a, b \in D$. Therefore $(a, b) \sim (a, b)\ \forall\ (a, b) \in S$.

Symmetry. We have $(a, b) \sim (c, d)$

$\Rightarrow ad = bc \Rightarrow da = cb$ [$\because$ Multiplication is commutative in D]

$\Rightarrow cb = da \Rightarrow (c, d) \sim (a, b)$.

Transitivity. Let $(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$.

Then $ad = bc$ and $cf = de$.

$\therefore$ $adf = bcf$ and $bcf = bde$.

$\therefore$ $adf = bde$

$\Rightarrow afd = bed$ [$\because$ D is a commutative ring]

$\Rightarrow afd - bed = 0 \Rightarrow (af - be)\, d = 0$

$\Rightarrow af - be = 0$ [$\because$ $d \neq 0$ and D is without zero divisors]

$\Rightarrow af = be \Rightarrow (a, b) \sim (e, f)$.

Thus $\sim$ is an equivalence relation in S. Therefore it will partition S into disjoint equivalence classes We shall denote the equivalence class containing (a, b) by $\frac{a}{b}$. Other notations to denote this equivalence class are $\overline{(a, b)}$ or $[a, b]$.

Then $\frac{a}{b} = \{(c, d) \in S : (c, d) \sim (a, b)\}$.

Obviously $\frac{a}{b} = \frac{c}{d}$ iff $(a, b) \sim (c, d)$ *i.e.*, iff $ad = bc$.

Also $\frac{a}{b} = \frac{ax}{bx}$ $\forall\ x \in D_0$. The reason is that

$(a, b) \sim (ax, bx)$ since $abx = bax$.

These equivalence classes are our quotients. Let F be the set of all such quotients *i.e.*, $F = \left\{\frac{a}{b} : (a, b) \in S\right\}$.

We now define addition and multiplication operations in F as follows :

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} \text{ and } \frac{a}{b}\frac{c}{d} = \frac{ac}{bd}.$$

Since D is without zero divisors, therefore $b \neq 0,\ d \neq 0 \Rightarrow bd \neq 0$.

Therefore both $\frac{ad+bc}{bd}$ and $\frac{ac}{bd}$ are elements of F. Thus F is closed with respect to addition and multiplication. We shall now show that both addition and multiplication in F are well defined. For this we are to show that if

$$\frac{a}{b}=\frac{a'}{b'} \text{ and } \frac{c}{d}=\frac{c'}{d'}, \text{ then } \frac{a}{b}+\frac{c}{d}=\frac{a'}{b'}+\frac{c'}{d'} \text{ and } \frac{a}{b}\,\frac{c}{d}=\frac{a'}{b'}\,\frac{c'}{d'}.$$

We have $\frac{a}{b}=\frac{a'}{b'} \Rightarrow ab'=ba'$ and $\frac{c}{d}=\frac{c'}{d'} \Rightarrow cd'=dc'$.

Now to show that $\frac{a}{b}+\frac{c}{d}=\frac{a'}{b'}+\frac{c'}{d'}$, we are to show that

$$\frac{ad+bc}{bd}=\frac{a'd'+b'c'}{b'd'} \text{ i.e., } (ad+bc)\,b'd'=bd\,(a'd'+b'c').$$

Now $(ad+bc)\,b'd'=adb'd'+bcb'd'=ab'dd'+bb'cd'$

$=ba'\,dd'+bb'\,dc'$ [∵ $ab'=ba'$ and $cd'=dc'$]

$=bda'd'+bdb'c'=bd\,(a'd'+b'c')$, which was desired.

Again to show that $\frac{a}{b}\,\frac{c}{d}=\frac{a'}{b'}\,\frac{c'}{d'}$ we are to show that

$$\frac{ac}{bd}=\frac{a'c'}{b'd'} \text{ i.e., } acb'd'=bda'c'.$$

Now $acb'd'=ab'cd'=ba'dc'=bda'c'$, which was desired.

Therefore both addition and multiplication are well defined on F. We shall now show that F is a field for these two operations.

Associativity of addition. We have

$$\left(\frac{a}{b}+\frac{c}{d}\right)+\frac{e}{f}=\frac{ad+bc}{bd}+\frac{e}{f}=\frac{(ad+bc)\,f+bde}{bdf}$$

$$=\frac{adf+bcf+bde}{bdf}=\frac{adf+b\,(cf+de)}{bdf}=\frac{a}{b}+\frac{cf+de}{df}=\frac{a}{b}+\left(\frac{c}{d}+\frac{e}{f}\right).$$

Commutativity of addition. We have

$$\frac{a}{b}+\frac{c}{d}=\frac{ad+bc}{bd}=\frac{cb+da}{db}=\frac{c}{d}+\frac{a}{b}.$$

Existence of additive identity. We have

$\frac{0}{a} \in F$ where $a \neq 0$. If $\frac{c}{d}$ is any element of F, then

$$\frac{0}{a}+\frac{c}{d}=\frac{0d+ac}{ad}=\frac{0+ac}{ad}=\frac{ac}{ad}=\frac{c}{d} \qquad [\because\ acd=adc].$$

∴ $\frac{0}{a}$ is the additive identity. It should be noted that

$\frac{0}{a}=\frac{0}{b}\ \forall\ a, b \in D_0$. Also $\frac{c}{d}=\frac{0}{a}$ iff $ca=d0$ *i.e.*, $c=0$.

Existence of additive inverse. If $\frac{a}{b} \in F$, then $\frac{-a}{b} \in F$.

Also we have, $\frac{-a}{b}+\frac{a}{b}=\frac{(-a)\,b+ba}{b^2}=\frac{0}{b^2}=\frac{0}{a}$ $[\because\ 0a=b^2 0]$

$\therefore$ $\frac{-a}{b}$ is the additive inverse of $\frac{a}{b}$.

Associativity of multiplication.

$$\left(\frac{a}{b}\,\frac{c}{d}\right)\frac{e}{f}=\frac{ac}{bd}\,\frac{e}{f}=\frac{(ac)\,e}{(bd)\,f}=\frac{a\,(ce)}{b\,(df)}=\frac{a}{b}\,\frac{ce}{df}=\frac{a}{b}\left(\frac{c}{d}\,\frac{e}{f}\right).$$

Commutativity of multiplication. We have

$$\frac{a}{b}\,\frac{c}{d}=\frac{ac}{bd}=\frac{ca}{db}=\frac{c}{d}\,\frac{a}{b}.$$

Existence of multiplicative identity. We have

$\frac{a}{a} \in F$ where $a \neq 0$. Also if $\frac{c}{d} \in F$, then

$\frac{a}{a}\,\frac{c}{d}=\frac{ac}{aa}=\frac{c}{d}$ $[\because\ (ac, ad) \sim (c, d)$ because $acd=adc]$.

$\therefore$ $\frac{a}{a}$ is the multiplicative identity. It should be noted that

$$\frac{a}{a}=\frac{b}{b}\ \forall\ a, b \in D_0.$$

Existence of multiplicative inverse of non-zero elements of *F*.

Let $\frac{a}{b}$ be any non-zero element of F. Then $a \neq 0$.

$\therefore$ $\frac{b}{a} \in F$. Also we have

$$\frac{a}{b}\,\frac{b}{a}=\frac{ab}{ba}=\frac{ab}{ab}=\frac{a}{a}=\text{the unity element.}$$

$\therefore$ $\frac{b}{a}$ is the multiplicative inverse of $\frac{a}{b}$.

Distributivity of multiplication over addition. We have

$$\frac{a}{b}\left(\frac{c}{d}+\frac{e}{f}\right)=\frac{a}{b}\,\frac{cf+de}{df}=\frac{a\,(cf+de)}{bdf}=\frac{acf+ade}{bdf}=\frac{(acf+ade)\,bdf}{bdf\,bdf}$$

$$=\frac{acfbdf+bdfade}{bdf\,bdf}=\frac{acf}{bdf}+\frac{ade}{bdf}=\frac{ac}{bd}+\frac{ae}{bf}=\frac{a}{b}\,\frac{c}{d}+\frac{a}{b}\,\frac{e}{f}.$$

Similarly the other distributive law holds.

$\therefore$ F is a field under the addition and multiplication as defined above. This field F is called the field of quotients of D.

We shall now show that the field F contains a subset D' such that D is isomorphic to D'.

Let $D'=\left\{\frac{ax}{x}\in F: a,\ 0\neq x\in D\right\}$. Then $D'\subseteq F$. If $x\neq 0$, $y\neq 0$ are elements of D, then $\frac{ax}{x}=\frac{ay}{y}$ since $axy=xay$. Therefore if x is any fixed non-zero element of D, we can write

$$D'=\left\{\frac{ax}{x}\in F: a\in D\right\}.$$

We claim that the function $\phi : D\to D'$ defined by

$\phi(a)=\frac{ax}{x}\ \forall\ a\in D$ is an isomorphism of D onto D'.

ϕ is one-one. We have

$$\phi(a)=\phi(b)\Rightarrow\frac{ax}{x}=\frac{bx}{x}\Rightarrow axx=xbx\Rightarrow ax^2=bx^2\Rightarrow(a-b)\,x^2=0$$
$$\Rightarrow a-b=0,\text{ since } x^2\neq 0$$
$$\Rightarrow a=b.$$

$\therefore$ ϕ is one-one.

ϕ is onto D'. If $\frac{ax}{x}\in D'$, then $a\in D$. Also we have $\phi(a)=\frac{ax}{x}$. Thus ϕ is onto D'.

Also $\phi(a+b)=\frac{(a+b)\,x}{x}=\frac{(a+b)\,x^2}{x^2}=\frac{ax^2+bx^2}{x^2}=\frac{axx+xbx}{x^2}$

$$=\frac{ax}{x}+\frac{bx}{x}=\phi(a)+\phi(b)$$

and $\phi(ab)=\frac{(ab)\,x}{x}=\frac{(ab)\,x^2}{x^2}=\frac{(ax)\,(bx)}{x^2}=\frac{ax}{x}\,\frac{bx}{x}=\phi(a)\,\phi(b)$.

$\therefore$ ϕ is an isomorphism of D onto D'.

Hence $D\cong D'$.

If we identify D' with D *i.e.,* if in F we write a, b, c etc. in place of $\frac{ax}{x}, \frac{bx}{x}, \frac{cx}{x}$ etc., then we see that D is contained in F. Thus F (the field of quotients of D) is a field containing D.

In the next theorem we shall show that the quotient field F of D is the smallest field containing D. In other words if D is contained in any other field K, then F will also be contained in K.

Theorem 2. *If K is any field which contains an integral domain D, then K contains a subfield isomorphic to the quotient field F of D.*

In other words the quotient field F of D is the smallest field containing D. **(Allahabad 1980; Banaras 69, 70)**

Proof. Let D be a commutative ring without zero divisors. Let $a \in D$ and $0 \neq b \in D$. Since K is a field containing D, therefore $a \in K$, $0 \neq b \in K \Rightarrow ab^{-1} \in K$.

Let K' be the subset of K containing the elements of the form ab^{-1} where $a, b \in D$ with $b \neq 0$. Thus

$$K' = \{ab^{-1} \in K : a, 0 \neq b \in D\}.$$

We shall show that K' is a subfield of K and K' is isomorphic to the quotient field F of D. Let $ab^{-1} \in K'$, $cd^{-1} \in K'$. Then

$$0 \neq b,\ 0 \neq d \in D.$$

Now $ab^{-1} - cd^{-1} = add^{-1}b^{-1} - cbb^{-1}d^{-1} = (ad - bc)\,d^{-1}b^{-1}$
$= (ad - bc)(bd)^{-1} \in K'$, since $ad - bc \in D$ and $0 \neq bd \in D$.

Further suppose that $0 \neq cd^{-1} \in K'$. Then $c \neq 0$ and we have $(ab^{-1})(cd^{-1})^{-1} = ab^{-1}dc^{-1} = ad\,(cb)^{-1} \in K'$, since $ad \in D$ and $0 \neq cb \in D$.

Hence K' is a subfield of K. We shall now show that the quotient field F of D is isomorphic to K'. We have

$$F = \left\{\frac{a}{b} : a \in D,\ 0 \neq b \in D\right\}.$$

Consider the mapping $f : F \to K'$ defined by

$$f\left(\frac{a}{b}\right) = ab^{-1}\ \forall\ \frac{a}{b} \in F.$$

The mapping f is one-one because we have

$$f\left(\frac{a}{b}\right) = f\left(\frac{c}{d}\right) \Rightarrow ab^{-1} = cd^{-1}$$
$$\Rightarrow ab^{-1}\,bd = cd^{-1}\,bd \Rightarrow ad = cb\,d^{-1}d$$
$$\Rightarrow ad = bc \Rightarrow (a, b) \sim (c, d) \Rightarrow \frac{a}{b} = \frac{c}{d}.$$

Also f is onto K'. If ab^{-1} is any element of K', then $\frac{a}{b} \in F$ and $f\left(\frac{a}{b}\right) = ab^{-1}$.

Further $f\left(\frac{a}{b} + \frac{c}{d}\right) = f\left(\frac{ad + bc}{bd}\right) = (ad + bc)(bd)^{-1}$
$$= (ad + bc)\,d^{-1}b^{-1} = add^{-1}\,b^{-1} + bc\,d^{-1}b^{-1}$$
$$= ab^{-1} + cd^{-1} = f\left(\frac{a}{b}\right) + f\left(\frac{c}{d}\right).$$

Also $f\left(\frac{a}{b}\frac{c}{d}\right) = f\left(\frac{ac}{bd}\right) = (ac)(bd)^{-1} = (ac)d^{-1}b^{-1}$

$$\Rightarrow (ab^{-1})\,(cd^{-1})=f\left(\frac{a}{b}\right)f\left(\frac{c}{d}\right).$$

Hence $\quad F\cong K'$.

If we identify K' with F, we see that if D is contained in any field K, then F is also contained in K. Therefore F is the smallest field containing D.

Cor. *The quotient field of a finite integral domain coincides with itself.*

Suppose D is a finite integral domain. Then D is also a field. Thus D is the smallest field containing D. The quotient field F of D is also the smallest field containing D. Hence F coincides with D.

Ex. *What is the quotient field of 2Z, where Z is the ring of integers ?* **(Poona 1973)**

Theorem. *Any two isomorphic integral domains have isomorphic quotient fields.* **(Banaras 1965)**

Proof. Suppose D and D' are two isomorphic integral domains.

Let f be an isomorphism of D onto D'. If a, b, c etc. are the elements of D then $f(a)$, $f(b)$, $f(c)$ etc. will be the elements of D'. Also

$$f(a+b)=f(a)+f(b) \text{ and } f(ab)=f(a)\,f(b)\ \forall\ a, b\in D.$$

Let F, F' be the quotient fields of D, D' respectively. Then F consists of the equivalence classes (quotients) of the form $\frac{a}{b}$ where $a, 0\neq b\in D$ and F' consists of the equivalence classes of the form $\frac{f(a)}{f(b)}$ where $f(a), 0\neq f(b)\in D'$.

Consider the mapping $\phi : F\to F'$ defined by

$$\phi\left(\frac{a}{b}\right)=\frac{f(a)}{f(b)}\ \forall\ \frac{a}{b}\in F.$$

First we shall show that the mapping ϕ is well defined *i.e.*, if $\frac{a}{b}=\frac{c}{d}$, then $\phi\left(\frac{a}{b}\right)=\phi\left(\frac{c}{d}\right)$. We have $\frac{a}{b}=\frac{c}{d}\Rightarrow ad=bc$

$$\Rightarrow f(ad)=f(bc)\Rightarrow f(a)\,f(d)=f(b)\,f(c)$$

$$\Rightarrow \frac{f(a)}{f(b)}=\frac{f(c)}{f(d)}\Rightarrow \phi\left(\frac{a}{b}\right)=\phi\left(\frac{c}{d}\right).$$

$\therefore\quad \phi$ is well-defined.

ϕ is one-one. We have

$$\phi\left(\frac{a}{b}\right)=\phi\left(\frac{c}{d}\right) \Rightarrow \frac{f(a)}{f(b)}=\frac{f(c)}{f(d)}$$
$$\Rightarrow f(a)\,f(d)=f(b)\,f(c) \Rightarrow f(ad)=f(bc)$$
$$\Rightarrow ad=bc \qquad [\because f \text{ is one-one}]$$
$$\Rightarrow \frac{a}{b}=\frac{c}{d}.$$

$\therefore$ ϕ is one-one.

Also ϕ is onto F'. If $\dfrac{f(a)}{f(b)} \in F'$, then $\dfrac{a}{b} \in F$ and $\phi\left(\dfrac{a}{b}\right)=\dfrac{f(a)}{f(b)}$. Therefore ϕ is onto F'.

Further $\phi\left(\dfrac{a}{b}+\dfrac{c}{d}\right)=\phi\left(\dfrac{ad+bc}{bd}\right)=\dfrac{f(ad+bc)}{f(bd)}=\dfrac{f(ad)+f(bc)}{f(b)\,f(d)}$

$$=\frac{f(a)\,f(d)+f(b)\,f(c)}{f(b)\,f(d)}=\frac{f(a)}{f(b)}+\frac{f(c)}{f(d)}=\phi\left(\frac{a}{b}\right)+\phi\left(\frac{c}{d}\right).$$

Also $\phi\left(\dfrac{a}{b}\,\dfrac{c}{d}\right)=\phi\left(\dfrac{ac}{bd}\right)=\dfrac{f(ac)}{f(bd)}=\dfrac{f(a)\,f(c)}{f(b)\,f(d)}$

$$=\frac{f(a)}{f(b)}\,\frac{f(c)}{f(d)}=\phi\left(\frac{a}{b}\right)\phi\left(\frac{c}{d}\right).$$

$\therefore$ ϕ is an isomorphism of F onto F'.

$\therefore$ $F \cong F'$.

§ 15. Ideals. Definition.

(a) Left Ideal. (Kurushetra 1970)

A non-empty subset S of a ring R is said to be a left ideal of R if :

(*i*) *S is a subgroup of R with respect to addition.*

(*ii*) $rs \in S \;\forall\; r \in R$ *and* $\forall\; s \in S$.

(b) Right Ideal. *A non-empty subset S of a ring R is said to be a right ideal of R if* :

(*i*) *S is a subgroup of R under addition.*

(*ii*) $sr \in S \;\forall\; r \in R$ *and* $\forall\; s \in S$.

(c) Ideal. **(Kurushetra 1970; Rajasthan 76; Kanpur 71; Nagarjuna 78, 79; Sambalpur 77; Meerut 81)**

A non-empty subset S of a ring R is said to be an *ideal* (also *a two sided ideal*) if and only if it is both a left ideal and a right ideal. *Thus a non-empty subset S of a ring R is said to be an ideal of R if* :

(*i*) *S is a subgroup of R under addition i.e., S is a subgroup of the additive group of R.*

(ii) $rs \in S$ *and* $sr \in S$ *for every* $r \in R$ *and for every* $s \in S$.

If S is an ideal of a ring R, then S is also a subring of R. The obvious reason is that S is a subgroup of R under addition and from condition (ii), we have $xs \in S \; \forall \; x, s \in S$ because $x \in S \Rightarrow x \in R$. Thus S is closed with respect to multiplication. Therefore S is a subring of R. *Thus every ideal of a ring R is also a subring of R.* But every subring is not an ideal. An ideal requires a stronger closure property than the subring. If S is a subgroup of R under addition, then S will be a subring if S is closed with respect to multiplication *i.e.*, the product of two elements of S is again in S. But S will be an ideal if the product of any element of S with any element of R is in S.

If R is a commutative ring, then every left ideal will also be a right ideal. Therefore in a commutative ring every left (right) ideal is an ideal.

Note 1. If we are to prove that a non-empty subset S of a ring R is an ideal of R, then it is sufficient to prove

(i) $a \in S, b \in S \Rightarrow a - b \in S$

and (ii) $rs \in S$ and $sr \in S \; \forall \; r \in R$ and $\forall \; s \in S$.

Obviously (i) is a sufficient condition for S to be a subgroup of R under addition.

If R is a commutative ring, then the condition (ii) will become more simple. Then it will become

$$rs \in S \; \forall \; r \in R \text{ and } \forall \; s \in S.$$

Note 2. Every ring R always possesses two improper ideals : one R itself and the other consisting of 0 only. These are respectively known as the **unit ideal** and the **null ideal**.

Any other ideals of R are called **proper ideals**. A ring having no proper ideals is called a **simple ring**.

Algebra of Ideals.

Theorem 1. *The intersection of any two left ideals of a ring is again a left ideal of the ring.* **(Meerut 1978)**

Proof. Let I_1 and I_2 be two left ideals of a ring R. Then I_1, I_2 are subgroups of R under addition. Therefore $I_1 \cap I_2$ is also a subgroup of R under addition.

Now to show that $I_1 \cap I_2$ is a left ideal of R, we are only to show that $r \in R, s \in I_1 \cap I_2 \Rightarrow rs \in I_1 \cap I_2$.

We have $s \in I_1 \cap I_2 \Rightarrow s \in I_1, s \in I_2$.

But I_1 and I_2 are left ideals of R. Therefore

$r \in R, s \in I_1 \Rightarrow rs \in I_1$ and $r \in R, s \in I_2 \Rightarrow rs \in I_2$.

Now $rs \in I_1, rs \in I_2 \Rightarrow rs \in I_1 \cap I_2$.

$\therefore$ $I_1 \cap I_2$ is also a left ideal of R.

Note. A similar result can be proved for right ideals as well as for ideals.

Theorem 2. *An arbitrary intersection of left ideals of a ring is a left ideal of the ring.* **(Banaras 1966)**

Proof. Let R be a ring and let $\{S_t : t \in T\}$ be any family of left ideals of R. Here T is an index set and is such that $\forall\ t \in T$, S_t is a left ideal of R. Let $S = \bigcap_{t \in T} S_t = \{x \in R : x \in S_t\ \forall\ t \in T\}$ be the intersection of this family of left ideals of R. Then to prove that S is also a left ideal of R.

Obviously $S \neq \varnothing$, since at least 0 is in $S_t\ \forall\ t \in T$.

Now let a, b be any two elements of S. Then

$$a, b \in S \Rightarrow a, b \in S_t\ \forall\ t \in T$$

$$\Rightarrow a - b \in S_t\ \forall\ t \in T$$

$$[\because\ \forall\ t \in T, S_t \text{ is a left ideal of } R]$$

$$\Rightarrow a - b \in \bigcap_{t \in T} S_t \Rightarrow a - b \in S.$$

Now let a be any element of S and r be any element of R.

We have $a \in S \Rightarrow a \in \bigcap_{t \in T} S_t \Rightarrow a \in S_t\ \forall\ t \in T$

$$\Rightarrow ra \in S_t\ \forall\ t \in T \qquad [\because\ \forall\ t \in T, S_t \text{ is a left ideal of } R]$$

$$\Rightarrow ra \in \bigcap_{t \in T} S_t \Rightarrow ra \in S.$$

Thus $a, b \in S \Rightarrow a - b \in S$ and $r \in R, a \in S \Rightarrow ra \in S$.

$\therefore$ S is a left ideal of R.

Smallest left ideal containing a given subset.

Definition. *Let M be a non-empty subset of a ring R. Then a left ideal I of R is called the smallest left ideal of R containing M, if I contains M and if I is contained in every left ideal of R containing M.*

The smallest left ideal of R containing M is called the **left ideal generated by** M and will be denoted by (M).

It can be easily seen that the intersection of the family of left ideals containing M is the left ideal generated by M.

Note. A similar definition can be given for the right ideal generated by M as well as for the ideal generated by M. For this purpose simply replace the word 'left ideal' by 'right ideal' or by 'ideal'.

Sum of two left ideals.

Theorem 3. *The left ideal generated by the union* $I_1 \cup I_2$ *of two left ideals is the set* I_1+I_2 *consisting of the elements of R obtained on adding any element of* I_1 *to any element of* I_2.

(Meerut 1970)

Proof. Let $a_1+a_2,\ b_1+b_2 \in I_1+I_2$.

Then $a_1, b_1 \in I_1$ and $a_2, b_2 \in I_2$.

Since I_1, I_2 are left ideals of R, therefore they are subgroups of the additive group of R. Therefore

$a_1, b_1 \in I_1 \Rightarrow a_1-b_1 \in I_1$ and $a_2, b_2 \in I_2 \Rightarrow a_2-b_2 \in I_2$.

Consequently $(a_1+a_2)-(b_1+b_2)=(a_1-b_1)+(a_2-b_2)\in I_1+I_2$.

Therefore I_1+I_2 is a subgroup of the additive group of R.

Now let $r \in R$ and $a_1+a_2 \in I_1+I_2$. Then $a_1 \in I_1$, $a_2 \in I_2$. We have $r\,(a_1+a_2)=ra_1+ra_2 \in I_1+I_2$.

[$\because$ I_1 is a left ideal implies $ra_1 \in I_1$ and similarly $ra_2 \in I_2$]

$\therefore$ I_1+I_2 is a left ideal of R.

Since $0 \in I_2$, therefore $a_1 \in I_1$ can be written as a_1+0. Thus

$$a_1 \in I_1 \Rightarrow a_1 \in I_1+I_2.$$

$\therefore \quad I_1 \subseteq I_1+I_2$.

Similarly $\quad I_2 \subseteq I_1+I_2$.

$\therefore \quad I_1 \cup I_2 \subseteq I_1+I_2$.

Thus I_1+I_2 is a left ideal containing $I_1 \cup I_2$.

Also if any left ideal contains $I_1 \cup I_2$, then it must contain I_1+I_2.

$\therefore$ I_1+I_2 is the smallest left ideal containing $I_1 \cup I_2$.

$\therefore$ $I_1+I_2=$ the left ideal generated by $I_1 \cup I_2=(I_1 \cup I_2)$.

Note. A similar result can be proved for right ideals as well as for ideals.

Solved Examples

Ex. 1. *The set N of all* 2×2 *matrices of the form*

$$\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix}$$

for a, b, integers is a left ideal but not a right ideal in the ring R of all 2×2 *matrices with elements as integers. Here N is the subset of R consisting of those elements whose second column contains only zeros.*

(Rajasthan 1977)

Solution. Let $A=\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix}$, $B=\begin{bmatrix} c & 0 \\ d & 0 \end{bmatrix}$ be any two elements of

N. Then $A-B=\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix}-\begin{bmatrix} c & 0 \\ d & 0 \end{bmatrix}=\begin{bmatrix} a-c & 0 \\ b-d & 0 \end{bmatrix} \in N$.

$\therefore$ N is a subgroup of R under addition.

Now let $U=\begin{bmatrix} w & x \\ y & z \end{bmatrix}$ be any element of R and $A=\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix}$ be any element of N.

Then $UA=\begin{bmatrix} w & x \\ y & z \end{bmatrix}\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix}=\begin{bmatrix} wa+xb & 0 \\ ya+zb & 0 \end{bmatrix} \in N$.

Therefore N is a left ideal of R. It is not a right ideal, since

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \in N, \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \in R,$$

and the product $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}=\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ which is not an element of N.

Ex. 2. *If m is a fixed integer, the set P of integers given by,*

$$P=\{xm : x \text{ is an integer}\}$$

is an ideal of the ring R of all integers.

Solution. Let x_1m and x_2m be any two elements of P. Then x_1 and x_2 are some integers.

We have $x_1m-x_2m=(x_1-x_2)\ m \in P$ since x_1-x_2 is also an integer.

$\therefore$ P is a subgroup of R under addition.

Now let r be any integer *i.e.*, r be any element of R and xm be any element of P. Then $r\,(xm)=(rx)\ m \in P$ since rx is also an integer. Therefore P is a left ideal of R. But R is a commutative ring. Hence P is an ideal of R.

Ex. 3. *The set of integers* **I** *is only a subring but not an ideal of the ring of rational numbers* $(\mathbf{Q},+,\cdot)$

(Meerut 1980, 81, 82; Rajasthan 77)

Solution. The product of a rational number and an integer is not necessarily an integer.

For example $3 \in \mathbf{I}$, $2/5 \in \mathbf{Q}$ but $(2/5).3=6/5 \notin \mathbf{I}$.

$\therefore$ **I** is not an ideal of the ring of rational numbers.

Ex. 4. *The set* **Q** *of rational numbers is only a subring but not an ideal of the ring of real numbers* $(\mathbf{R},+,\cdot)$. **(Meerut 1983 P)**

Solution. The product of a rational number and a real number is not necessarily a real number. For example $\frac{1}{2} \in \mathbf{Q}$, $\sqrt{7} \in \mathbf{R}$ but $\frac{1}{2}.\sqrt{7}=\frac{\sqrt{7}}{2} \notin \mathbf{Q}$.

$\therefore$ **Q** is not an ideal of the ring of real numbers.

Ex. 5. *Prove that the subset S of all matrices of the form*

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

with a and b integers, forms a subring of the ring R of all 2×2 matrices having elements as integers. Prove further that S is neither a right ideal nor a left in R.

Solution. Let $A=\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, $B=\begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix}$ be any two elements of S. Then $A-B=\begin{bmatrix} a-c & 0 \\ 0 & b-d \end{bmatrix} \in S$.

Also $AB=\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}\begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix}=\begin{bmatrix} ac & 0 \\ 0 & bd \end{bmatrix} \in S$.

$\therefore$ S is a subring of R.

Further $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in S$, $\begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} \in R$ and the product $\begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}=\begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} \notin S$. Therefore S is not a left ideal.

Again $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix}=\begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} \notin S$. Therefore S is not a right ideal.

Ex. 6. *If U is an ideal of a ring R with unity and* $1 \in U$ *prove that* $U=R$. **(Sambalpur 1977)**

Solution. We have $U \subseteq R$ since U is an ideal of R. Let x be any element of R. Since U is an ideal of R, therefore

$$1 \in U, x \in R \Rightarrow 1x \in U \Rightarrow x \in U.$$

$$\therefore \quad R \subseteq U.$$

$$\therefore \quad U=R.$$

Ex. 7. *If R is a ring and* $a \in R$ *let* $T=\{x \in R : ax=0\}$. *Prove that T is a right ideal of R.*

Solution. First we see that T is not empty because

$$0 \in R \text{ is such that } a0=0.$$

Let x_1, x_2 be any two elements of T. Then $ax_1=0$, $ax_2=0$.

We have $a(x_1-x_2)=ax_1-ax_2=0-0=0$.

$\therefore$ $x_1-x_2 \in T$.

$\therefore$ T is a subgroup of R under addition.

Now to show that T is a right ideal of R we are to show that $x \in T, y \in R \Rightarrow xy \in T$. But $x \in T \Rightarrow ax=0$. If we show that $a(xy)=0$, then xy will be an element of T.

We have $a(xy)=(ax)y=0y=0$.

$\therefore \quad xy \in T.$

$\therefore$ T is a right ideal of R.

Ex 8. *Prove that the intersection of two ideals of R is an ideal of R.* **(Vikram 1976; Kumayon 78; Meerut 80, 81, 83, 84 P)**

Solution. Let S and T be two ideals of R. Then S, T are subgroups of R under addition. Therefore $S \cap T$ is also a subgroup of R under addition.

Now to show that $S \cap T$ is an ideal of R, we are only to show that

$$r \in R,\ s \in S \cap T \Rightarrow rs \in S \cap T,\ sr \in S \cap T.$$

We have $s \in S \cap T \Rightarrow s \in S,\ s \in T.$

But S and T are ideals of R. Therefore

$$r \in R,\ s \in S \Rightarrow rs \in S,\ sr \in S$$

and $\quad r \in R,\ s \in T \Rightarrow rs \in T,\ sr \in T.$

Now $rs \in S,\ rs \in T \Rightarrow rs \in S \cap T$

and $\quad sr \in S,\ sr \in T \Rightarrow sr \in S \cap T.$

$\therefore$ $S \cap T$ is also an ideal of R.

Ex. 9. *Show that S is an ideal of $S+T$ where S is any ideal of ring R, and T any subring of R.* **(Meerut 1984)**

Solution. Since S is an ideal of R therefore S is a subring of R. Also T is a subring of R. First we shall show that $S+T$ is a subring of R. Let $a+\alpha$, $b+\beta \in S+T$, where $a, b \in S$ and $\alpha, \beta \in T$.

Since S is a subring, therefore $a-b \in S$. Similarly $\alpha-\beta \in T$.

$\therefore \quad (a+\alpha)-(b+\beta)=(a-b)+(\alpha-\beta) \in S+T.$

Also $(a+\alpha)(b+\beta)=ab+a\beta+\alpha b+\alpha\beta=(ab+a\beta+\alpha b)+\alpha\beta.$

Now S is a subring. Therefore $a, b \in S \Rightarrow ab \in S$.

Also S is an ideal, therefore $a, b \in S$ and $\alpha, \beta \in R \Rightarrow a\beta$, $\alpha b \in S$. Therefore $ab+a\beta+\alpha b \in S$.

Further T is a subring implies $\alpha\beta \in T$ if $\alpha, \beta \in T$.

$\therefore \quad (a+\alpha)(b+\beta)=(ab+a\beta+\alpha b)+\alpha\beta \in S+T.$

$\therefore \quad S+T$ is a subring of R.

Since $0 \in T$, therefore $a \in S$ can be written as

$$a=a+0 \in S+T.$$

$$\therefore \quad S \subseteq S+T.$$

Thus $S \subseteq S+T$ and $S+T$ is a subring of R. Since S is an ideal of R, therefore S is also an ideal of $S+T$.

Ex. 10. *If U is a left ideal of a ring R, let*

$$\lambda(U)=\{x \in R : xu=0\ \forall\ u \in U\}.$$

Prove that $\lambda(U)$ is a two sided ideal of R.

Solution. First we see that $\lambda(U) \neq \emptyset$ because $0 \in R$ is such that
$$0u=0 \ \forall \ u \in U.$$

Now let x_1, x_2 be any two elements of $\lambda(U)$. Then
$$x_1u=0 \ \forall \ u \in U \text{ and } x_2u=0 \ \forall \ u \in U.$$

We have $(x_1-x_2)\, u=x_1u-x_2u=0-0=0$ for all $u \in U$.

$\therefore \quad x_1-x_2 \in \lambda(U)$.

Now let x be any element of $\lambda(U)$ and r be any element of R.

Then $\quad xu=0 \ \forall \ u \in U \quad$ [by def. of $\lambda(U)$]

$\Rightarrow r(xu)=r0 \ \forall \ u \in U$

$\Rightarrow (rx)\, u=0$ for all $u \in U \Rightarrow rx \in \lambda(U)$.

Further U is a left ideal of R. Therefore $ru \in U \ \forall \ u \in U$.

Since $x \in \lambda(U)$, therefore by def. of $\lambda(U)$, we have

$x \in \lambda(U)$, $ru \in U \Rightarrow x(ru)=0$ for all $u \in U$

$\Rightarrow (xr)\, u=0$ for all $u \in U$

$\Rightarrow xr \in \lambda(U)$.

Thus $x \in \lambda(U)$, $r \in R \Rightarrow xr, rx \in \lambda(U)$.

Hence $\lambda(U)$ is a two sided ideal of R.

Ex. 11. *For any given element a of a ring R let*
$$Ra=\{xa : x \in R\}.$$
Prove that Ra is a left ideal of R.

Solution Let x_1a, x_2a be any two elements of Ra where $x_1, x_2 \in R$. We have $x_1a-x_2a=(x_1-x_2)\, a \in Ra$, since
$$x_1, x_2 \in R \Rightarrow x_1-x_2 \in R.$$

Thus $x_1a, x_2a \in Ra \Rightarrow x_1a-x_2a \in Ra$.

Now let xa be any element of Ra where $x \in R$ and r be any element of R. We have $r(xa)=(rx)\, a \in Ra$, since
$$r \in R, x \in R \Rightarrow rx \in R.$$

Thus $r \in R$, $xa \in Ra \Rightarrow r(xa) \in Ra$.

$\therefore \quad Ra$ is a left ideal of R.

Ex. 12. *If U is an ideal of a ring R, let*
$$[R:U]=\{x \in R : rx \in U \ \forall \ r \in R\}.$$
Prove that $[R:U]$ is an ideal of R and that it contains U.

Solution. First we see that $[R:U]$ is not empty because $0 \in R$ is such that $r0=0 \in U$ for all $r \in R$.

Now let x_1, x_2 be any two elements of $[R:U]$. Then $rx_1 \in U \ \forall \ r \in R$, and $rx_2 \in U \ \forall \ r \in R$.

Since U is an ideal, therefore
$$rx_1 \in U, rx_2 \in U \Rightarrow rx_1-rx_2 \in U$$

$\Rightarrow r(x_1-x_2) \in U \; \forall \; r \in R$

$\Rightarrow x_1-x_2 \in [R:U]$, by def. of $[R:U]$.

Now let x be any element of $[R:U]$ and s be any element of R. Then $rx \in U \; \forall \; r \in R$

$\Rightarrow (rx)\, s \in U \; \forall \; r \in R$ [$\because$ U is an ideal and so $s \in R$, $rx \in U \Rightarrow (rx)\, s \in U$]

$\Rightarrow r\,(xs) \in U$ for all $r \in R \Rightarrow xs \in [R:U]$.

Also $rx \in U \; \forall \; r \in R \Rightarrow sx \in U$ [$\because$ $s \in R$]

$\Rightarrow (sx)\, r \in U \; \forall \; r \in R$

[$\because$ U is an ideal and so $sx \in U$, $r \in R \Rightarrow (sx)\, r \in U$]

$\Rightarrow sx \in [R:U]$.

Thus $x \in [R:U]$, $s \in R \Rightarrow xs \in [R:U]$, $sx \in [R:U]$.

Therefore $[R:U]$ is an ideal of R.

Now to show that $U \subseteq [R:U]$. We have

$y \in U \Rightarrow yr \in U \; \forall \; r \in R$ [$\because$ U is an ideal]

$\Rightarrow y \in [R:U]$.

$\therefore \; U \subseteq [R:U]$.

Ex. 13. *If U, V are ideals of a ring R let UV be the set of all those elements of R which can be written as finite sums of elements of the form uv where $u \in U$ and $v \in V$. Prove that UV is an ideal of R.*

(Marathwada 1972; Lucknow 70; Meerut 70)

Solution. U and V are ideals of a ring R. Let

$UV = \{u_1v_1+u_2v_2+\ldots+u_nv_n : u_1, u_2, \ldots, u_n \in U, v_1, v_2, \ldots, v_n \in V$ and n is any positive integer$\}$.

To prove that UV is also an ideal of R.

Let $\alpha = u_1v_1+u_2v_2+\ldots+u_nv_n$, $\beta = u_1'v_1'+u_2'v_2'+\ldots+u_m'v_m'$ be any two elements of UV, where $u_1, u_2, \ldots, u_n, u_1', u_2', \ldots, u_m' \in U$ and $v_1, v_2, \ldots, v_n, v_1', v_2', \ldots, v_m' \in V$. Also m and n are any positive integers.

We have $\alpha-\beta = u_1v_1+u_2v_2+\ldots+u_nv_n-u_1'v_1'-u_2'v_2'-\ldots-u_m'v_m'$
$= u_1v_1+u_2v_2+\ldots+u_nv_n+(-u_1')\,v_1'+(-u_2')\,v_2'+\ldots+(-u_m')\,v_m'$.

This is obviously an element of UV because U is an ideal and therefore $u_1' \in U \Rightarrow (-u_1') \in U$, etc.

Again let $r \in R$ and $\alpha \in UV$. Then

$r\alpha = r\,(u_1v_1+u_2v_2+\ldots+u_nv_n) = (ru_1)\,v_1+(ru_2)\,v_2+\ldots+(ru_n)\,v_n$.

This is an element of UV because U is an ideal and therefore $r \in R$, $u_1 \in U \Rightarrow ru_1 \in U$, etc.

Also $\alpha r=(u_1v_1+u_2v_2+\ldots+u_nv_n)\ r=u_1\ (v_1r)+u_2\ (v_2r)+\ldots+u_n\ (v_nr)$.

This is an element of UV because V is an ideal and therefore $r\in R,\ v_1\in V \Rightarrow v_1r\in V$, etc.

Hence UV is an ideal of R.

Now to show that $UV \subseteq U \cap V$.

Let $\alpha=u_1v_1+\ldots+u_nv_n$ be any element of UV where

$$u_1,\ldots,\ u_n\in U \text{ and } v_1,\ldots,\ v_n\in V.$$

Now $v_1\in V \Rightarrow v_1\in R$. Also U is an ideal. Therefore

$$v_1\in R,\ u_1\in U \Rightarrow u_1v_1\in U.$$

Similarly $u_1\in U \Rightarrow u_1\in R$. But V is an ideal. Therefore

$$u_1\in R,\ v_1\in V \Rightarrow u_1v_1\in V.$$

Thus $u_1v_1\in U,\ u_1v_1\in V \Rightarrow u_1v_1\in U\cap V$.

Similarly $u_2v_2,\ldots,\ u_nv_n\in U\cap V$.

Since $U\cap V$ is also an ideal of R, therefore

$u_1v_1,\ldots,\ u_nv_n\in U\cap V \Rightarrow \alpha=u_1v_1+\ldots+u_nv_n\in U\cap V$.

Thus $\alpha\in UV \Rightarrow \alpha\in U\cap V$. Therefore $UV \subseteq U\cap V$.

Exercises

1. Distinguish between Subrings and Ideals in a ring. Show that the 2-rowed matrices of the form

$$\begin{bmatrix} a & 0 \\ b & c \end{bmatrix}$$

where a, b, c are integers form a subring of the ring of all 2-rowed matrices with integral entries. Is this subring an integral domain?

Ans. No. **(Kerala 1970; Meerut 73, 75, 77, 78, 79)**

2. Show that the set M of all 2×2 matrices of the form

$$\begin{bmatrix} 0 & a \\ 0 & b \end{bmatrix}$$

a, b integers is a left ideal but not a right ideal in the ring of all 2×2 matrices with elements as integers. **(Meerut 1970)**

3. Show that for a field F, the set of all matrices of the form

$$\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$$

for $a,\ b\in F$ is a right ideal but not a left ideal of the ring of all 2×2 matrices over the field F. **(I.A.S. 1970; Delhi 70)**

4. If U, V are ideals of a ring R, let

$$U+V=\{u+v : u\in U,\ v\in V\}.$$

Prove that $U+V$ is also an ideal of R.

(Marathwada 1972; Meerut 81)

5. Show that an arbitrary intersection of ideals of a ring is an ideal of the ring. **(Banaras 1961)**

6. If U is an ideal of a ring R, let

$$r(U)=\{x \in R : xu=0 \text{ for all } u \in U\}.$$

Prove that $r(U)$ is an ideal of R. **(Guru Nanak 1982)**

7. Consider the ring R of all 3×3 matrices of the type

$$\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix},$$

a, b, c, d, e, f are real numbers. Show that the set I of all matrices of the form

$$\begin{bmatrix} a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is a left ideal of R, which is not a right ideal. **(Meerut 1974)**

8. Verify the following for being true or false :

(i) The set of all positive rationals is a subring of the ring of all rational numbers.

(ii) A subring of any field is a field.

(iii) Any subring of the ring of integers, Z, is an ideal of Z. **(Meerut 1976)**

Ans. (i) and (ii) are false; (iii) is true.

§ 16. More about ideals.

Theorem 1. *A field has no proper ideals i.e., if F is a field then its only ideals are* (0) *and F itself.*

(I.A.S. 1972; Poona 73; Meerut 76, 80; Delhi 70; Andhra 77; Nagarjuna 80; Banaras 70; Sambalpur 77)

Proof. Let S be any non-zero ideal of the field F and let a be any non-zero element of S. We have $a^{-1} \in F$.

Since S is an ideal, therefore

$$a \in S,\ a^{-1} \in F \Rightarrow aa^{-1} \in S \Rightarrow 1 \in S.$$

Now let x be any element of F. Then

$$1 \in S,\ x \in F \Rightarrow 1x \in S \Rightarrow x \in S.$$

Thus each element of F belongs to S. Therefore $F \subseteq S$. But $S \subseteq F$. Therefore $S=F$.

Thus the only ideals of F are (0) and F itself.

Theorem 2. *If R is a commutative ring and $a \in R$, then*

$Ra=\{ra : r \in R\}$ is an ideal of R.

(Kanpur 1980: Madras 77)

Proof. In order to prove that Ra is an ideal of R, we should prove that Ra is a subgroup of R under addition and that if $u \in Ra$ and $x \in R$ then xu and ux are also in Ra. But R is a commutative ring, therefore $xu = ux$. Thus we only need to check that xu is in Ra.

Now, let $u, v \in Ra$. Then $u = r_1 a$, $v = r_2 a$ for some $r_1, r_2 \in R$.

We have $u - v = r_1 a - r_2 a = (r_1 - r_2)\, a \in Ra$ since $r_1 - r_2 \in R$.

Thus $u, v \in Ra \Rightarrow u - v \in Ra$. Hence Ra is a subgroup of R under addition.

Now let $x \in R$.

Then $xu = x\,(r_1 a) = (xr_1) a \in Ra$ since $xr_1 \in R$.

$\therefore$ Ra is an ideal of R.

Theorem 3. *A commutative ring with unity is a field if it has no proper ideals.*

(I A.S. 1973; Guru Nanak 82; Nagarjuna 79, 80; Meerut 76, 82 P; Kanpur 71; Vikram 76; Andhra 77; Rajasthan 77)

Proof. Let R be a commutative ring with unity having no proper ideals *i.e.*, the only ideals of R are (0) and R itself. In order to show that R is a field, we should show that each non-zero element of R possesses multiplicative inverse.

Let a be any non-zero element of R.

The set $Ra = \{ra : r \in R\}$ is an ideal of R. [See theorem 2]

Since $1 \in R$, therefore $1a = a \in Ra$. Thus $0 \neq a \in Ra$. Therefore the ideal $Ra \neq (0)$. Since R has no proper ideals, therefore the only possibility is that $Ra = R$. Thus every element of R is a multiple of a by some element of R. In particular, $1 \in R$ so it can be realised as a multiple of a. Thus there exists an element $b \in R$ such that $ba = 1$. Therefore $a^{-1} = b$. Hence each non-zero element of R possesses multiplicative inverse.

$\therefore$ R is a field

Ex *Prove that a commutative ring R with identity is a field if and only if it has no proper ideals.* **(Madras 1978)**

Theorem 4. *Let R be a ring with unit element, R not necessarily commutative, such that the only right ideals of R are* (0) *and R. Prove that R is division ring.* **(Kanpur 1971; Meerut 73)**

Proof. Let R be a ring with unity element having no proper right ideals *i.e.*, the only right ideals of R are (0) and R itself. In order to show that R is a division ring, we should show that each non-zero element of R possesses multiplicative inverse.

Let a be any non-zero element of R. Then $aR=\{ar : r \in R\}$ is a right ideal of R. Since $1 \in R$, therefore $a1=a \in aR$. Thus $0 \neq a \in aR$. Therefore the right ideal $aR \neq (0)$. Since R has no proper right ideals therefore the only possibility is that $aR=R$.

Now $1 \in R \Rightarrow 1 \in aR$ $\quad [\because \; aR=R]$

$\Rightarrow 1=ab$ for some $b \in R$

$\Rightarrow b$ is the right multiplicative inverse of a.

Thus each non-zero element of R possesses right multiplicative inverse.

Note that $1=ab \Rightarrow b \neq 0$ because if $b=0$, then

$$ab=a0=0 \neq 1.$$

Now it remains to show that b is also the left multiplicative inverse of a.

We have $(ba)\, b=b\,(ab)=b1=b$.

Since b is also a non-zero element of R, therefore b also possesses right multiplicative inverse. Let $b^{-1} \in R$ be such that $bb^{-1}=1$.

We have $(ba)\, b=b \Rightarrow (ba)\, bb^{-1}=bb^{-1} \Rightarrow (ba)\, 1=1 \Rightarrow ba=1$.

Thus b is also the left inverse of a and so b is the inverse of a.

Hence each non-zero element of R is inversible.

$\therefore$ R is a division ring.

Theorem 5. *Let R be a ring with unity element such that the only left ideals of R are (0) and R. Show that R is a division ring.*

(Sagar 1967; Kurushetra 70)

Proof. Proceed exactly as in the theorem 4. Here note that $Ra=\{ra : r \in R\}$ is a left ideal of R.

§ 17. **Ideal generated by a given subset of a ring.** If M is any subset of a ring R, we can find ideals containing M. For example, the ring R itself is an ideal containing any subset of R.

Smallest ideal containing a subset. Let M be any arbitrary subset of a ring R. Then an ideal S of R is called the *smallest ideal of R containing M if*

$$M \subseteq S,$$

and if S is contained in every ideal of R containing M.

Definition. *Let R be a ring and let M be an arbitrary subset of R. The smallest ideal of R containing M is said to be the ideal generated by M and is denoted by (M).*

In particular, if M consists of a single element, say a, of the ring R we write (a) in place of M. An ideal such as (a) generated by a single element of the ring is called a *principal ideal.*

Principal ideal. Definition. *An ideal S of a ring R is said to be a principal ideal if there exists an element $a \in S$ such that any ideal T of R containing a also contains S i.e., $S=(a)$.*

(Lucknow 1970)

Thus an ideal generated by a single element of itself is called a principal ideal.

If a ring R has a unity element 1, then the ideal generated by 1 is the whole ring *i.e.*, $(1)=R$, since every element $r \in R$ may be written as $r1$. For this reason ring itself is called the **unit ideal.** The ideal generated by the zero element of R *i.e.*, (0) consists of the zero element alone and is called the **null ideal.** . Every ring R has at least one principal ideal, namely, (0). Every ring with unity has at least two principal ideals, (0) and (1).

Theorem 1. *If a is an element in a commutative ring R with unity, then the set $S=\{ra : r \in R\}$ is a principal ideal of R generated by the element a i.e., $S=(a)$.*

Proof. First we should prove that $a \in S$. Since R is a ring with unit element 1, therefore $1a=a \in S$.

Now we should prove that S is an ideal of R. So first we should prove that S is a subgroup of R under addition. Let u, v be any two elements of S. Then $u=r_1a$, $v=r_2a$ for some $r_1, r_2 \in R$.

We have $u-v=r_1a-r_2a=(r_1-r_2)\,a \in S$ since $r_1-r_2 \in R$.

$\therefore$ S is subgroup of R under addition.

Now we should prove that $x \in R, u \in S \Rightarrow xu \in S$ and $ux \in S$. But R is a commutative ring, therefore $xu=ux$ and thus it is sufficient to show that $xu \in S$.

We have $xu=x\,(r_1a)=(xr_1)\,a \in S$ since $xr_1 \in R$.

$\therefore$ S is an ideal of R and $a \in S$.

Now in order to prove that S is an ideal generated by the element a, we should prove that if T is an ideal of R and $a \in T$, then $S \subseteq T$.

Let ra be any element of S. Then $r \in R$. If T is an ideal of R containing a, then $a \in T$, $r \in R \Rightarrow ra \in T$. Thus $S \subseteq T$.

Hence S is a principal ideal of R generated by the element a.

Example. *Suppose we are to find the principal ideal generated by 5 in the ring of integers.* The ring I of integers is a commutative ring with unity. Therefore $(5)=\{5r : r \in I\}$.

Thus the principal ideal generated by 5 is given by

$(5)=\{\ldots, -10, -5, 0, 5, 10, \ldots\}$.

Obviously $(-5)=(5)$.

Theorem 2. *Let S be an ideal of a commutative ring R. Let a be an element of S such that*

$$x \in S \Rightarrow x=ya \text{ for some } y \in R.$$

Then S is a principal ideal of R generated by a.

Proof. As given in the statement of the theorem, S is an ideal of R containing the element a. Let T be any ideal of R containing a. Then S will be a principal ideal of R generated by a if $S \subseteq T$.

Let x be any element of S. Then $x=ya$ for some $y \in R$.

Now $\quad y \in R, a \in T \Rightarrow ya \in T \quad$ [$\because$ T is an ideal]

$\Rightarrow x \in T. \quad$ [$\because$ $x=ya$]

Thus $\quad x \in S \Rightarrow x \in T.$

$\therefore \quad S \subseteq T.$

Hence S is a principal ideal of R generated by a.

Note. The above theorem will be very helpful in proving that an ideal S of a commutative ring R is a principal ideal. If we are able to find an element a in S such that

$$x \in S \Rightarrow x=ay \text{ for some } y \in R,$$

then S will be a principal ideal of R generated by a.

§ 18 Principal Ideal Ring. Definition.

(Meerut 1981, 84 P; Rajasthan 78; Guru Nanak 82; Andhra 77; Madras 83)

A commutative ring R without zero divisors and with unity element is a principal ideal ring if every ideal S in R is a principal ideal i.e., if every ideal S in R is of the form $S=(a)$ for some $a \in S$.

Theorem 1. *The ring of integers is a principal ideal ring.*

(Banaras 1971; Lucknow 70; Rajasthan 77; Andhra 77; Meerut 84 P; Nagarjuna 78; Madras 83)

Proof. Let $(\mathbf{I}, +, \cdot)$ be the ring of integers. Obviously **I** is a commutative ring with unity and without zero divisors. Therefore **I** will be a principal ideal ring if every ideal in **I** is a principal ideal.

Let S be any ideal of the ring of integers. If S is the null ideal, then $S=(0)$ so that S is a principal ideal.

So let us suppose that $S \neq (0)$.

Now S contains at least one non-zero integer, say a. Since S is a subgroup of R under addition, therefore $a \in S \Rightarrow -a \in S$. This shows that S contains at least one positive integer because if $0 \neq a$, then one of a and $-a$ must be positive.

Let S_+ be the set of all positive integers in S. Since S_+ is not empty, therefore by the well ordering principle S_+ must possess a least positive integer. Let s be this least element. We will now show that S is the principal ideal generated by s *i.e*, $S=(s)$.

Suppose now that n is any integer in S. Then by division algorithm, there exist integers q and r such that $n=qs+r$ with $0 \leqslant r < s$.

Now $\quad s \in S, q \in \mathbf{I} \Rightarrow qs \in S \qquad [\because \; S$ is an ideal$]$

and $\quad n \in S, qs \in S \Rightarrow n-qs \in S \quad [\because \; S$ is a subgroup of the additive group of $\mathbf{I}]$

$\Rightarrow r \in S. \qquad [\because \; n-qs=r]$

But $0 \leqslant r < s$ and s is the least positive integer such that $s \in S$. Hence r must be 0.

$$n=qs.$$

Thus $n \in S \Rightarrow n=qs$ for some $q \in \mathbf{I}$

Hence S is a principal ideal of $\mathbf{I}$ generated by s.

Since S was an arbitrary ideal in the ring of integers, therefore the ring of integers is a principal ideal ring.

Theorem 2 *Every field is a principal ideal ring.*

Proof A field has no proper ideals. The only ideals of a field are (i) the null ideal which is a principal ideal generated by 0 and (ii) the field itself which is also a principal ideal generated by 1. Thus a field is always a principal ideal ring.

§ 19. Divisibility in an Integral Domain. Definition. *Suppose* $0 \neq a$ *is an element of a commutative ring* R. *Then* a *is said to divide* $b \in R$, *if there exists an element* $c \in R$ *such that* $b=ac$.

We shall use the symbol $a \mid b$ to represent the fact that a divides b. Also if a divides b then sometimes we say that a is a factor of b or b is divisible by a or a is a divisor of b. From the definition of divisibility it follows that every non-zero element of R is a divisor of its zero element. Obviously we can write $0=a0$. Therefore if $0 \neq a \in R$, then $a \mid 0$.

Example 1. In the ring $\mathbf{I}$ of integers, we have $3 \mid 6$ since we have $6=3 \times 2$ and $2 \in \mathbf{I}$.

However in the ring of integers 3 is not a divisor of 7.

Example 2. In the ring $\mathbf{Q}$ of rational numbers, we have $3 \mid 7$ since we have $7=3 \times (7/3)$ and $7/3 \in \mathbf{Q}$.

Theorem 1. *If* R *is a commutative ring, then* *(i)* $a \mid b$ *and* $b \mid c \Rightarrow a \mid c$ *i.e., the relation of divisibility in* R *is a transitive relation.* **(Allahabad 1957)**

(ii) $a \mid b$ *and* $a \mid c \Rightarrow a \mid (b+c)$.

(iii) $a \mid b \Rightarrow a \mid bx$ *for all* $x \in R$.

Proof. (i) $a \mid b \Rightarrow b = ap$ for some $p \in R$

and $b \mid c \Rightarrow c = bq$ for some $q \in R$.

Now $c = bq$ and $b = ap \Rightarrow c = (ap)\, q \Rightarrow c = a\,(pq)$

$\Rightarrow a \mid c$ since $pq \in R$.

(ii) $a \mid b \Rightarrow b = ap$ for some $p \in R$

and $a \mid c \Rightarrow c = aq$ for some $q \in R$

Now $b = ap$ and $c = aq \Rightarrow b + c = ap + aq \Rightarrow b + c = a\,(p+q)$

$\Rightarrow a \mid (b+c)$ since $(p+q) \in R$.

(iii) $a \mid b \Rightarrow b = ap$ for some $p \in R$.

Now $b = ap \Rightarrow bx = (ap)\, x \;\forall\; x \in R$

$\Rightarrow bx = a\,(px) \Rightarrow a \mid bx$ since $px \in R$.

Units. (Raj. 77). *Let* R *be a commutative ring with unity element 1. An element* $a \in R$ *is a unit in* R *if there exists an element* $b \in R$ *such that* $ab = 1$. In other words units of R are those elements of R which possess multiplicative inverse.

The students should not confuse a unit with the unit element or the unity element of the ring. There may be more than one units in a ring but the unity element is always unique. Of course the unity element is also one of the units.

In the ring of integers I, the only units are 1 and -1. These are the only inversible elements of the ring of integers.

Every non-zero element of a field possesses multiplicative inverse. Therefore every non-zero element of a field is a unit.

It is obvious that if a is a unit in a ring R, then a^{-1} is also a unit in R. Also the product of two units is again a unit. Because if a, b are two units in R, then $(ab)^{-1} = b^{-1} a^{-1}$ which is an element of R. Of course *the set of all units in* R *forms a group under multiplication.* **(Madras 1975)**

Example. *Find all the units of the integral domain of Gaussian integers.*

Solution. Let $D = \{a + ib : a, b \in \text{I the set of integers}\}$ be the ring of Gaussian integers. The element $1 + 0i$ is the unity element of this ring. Let $x + iy$ be a unit and $x' + iy'$ be its inverse.

Then $(x+iy)\,(x'+iy') = 1 + 0i$

or $(xx' - yy') + i(xy' + yx') = 1 + 0i.$

Equating real and imaginary parts, we get

$$xx'-yy'=1 \qquad \ldots(1)$$

and
$$xy'+yx'=0. \qquad \ldots(2)$$

Squaring and adding (1) and (2), we get

$$x^2x'^2+y^2y'^2+x^2y'^2+y^2x'^2=1$$

or
$$(x^2+y^2)(x'^2+y'^2)=1.$$

Now the product of two positive integers can be equal to 1 if and only if each of them is 1.

$$\therefore \quad x^2+y^2=1.$$

This gives $x^2=0,\ y^2=1$

or $x^2=1,\ y^2=0.$

Thus $x=0,\ y=\pm 1$

or $x=\pm 1,\ y=0.$

$\therefore$ The only units of the integral domain of Gaussian integers are $0\pm i$, $(\pm 1)+0i$ *i.e.*, $1, -1, i, -i$.

Associates. Definition. (**Raj. 1977**). *Let R be a commutative ring with unity element 1. Then an element a of R is said to be an associate of* $b \in R$ *if* $a=ub$ *for some unit u in R.*

In symbols, we express it by writing $a \sim b$ which is read as '*a* is an associate of *b*' or '*a* and *b* are associates'.

From this definition we observe that in a commutative ring with unity all the associates of an element can be obtained by multiplying the element by different units in that ring

Illustrations.

1. The only units of the integral domain of integers are 1 and -1. Therefore if a is any non-zero integer, then it has exactly two associates namely $1a$ and $(-1)a$ *i.e.*, a and $-a$. Thus the two associates of 5 are 5 and -5.

2. In any commutative ring with unity the associate of 0 is only zero.

3. The only units of the domain of Gaussian integers are $1, -1, i, -i$. Therefore if $a+ib$ is any non-zero element of this domain, then it has exactly four associates namely,

$$1\,(a+ib),\ -1\,(a+ib),\ i\,(a+ib),\ -i\,(a+ib)$$

i.e., $a+ib,\ -a-ib,\ -b+ia,\ b-ia.$

Theorem 2. *Let R be a commutative ring with unity element 1. The relation in R defined by 'a is an associate of b' is an equivalence relation.*

Proof.

Reflexivity. Let a be any element of R. Then $a=1a$. Therefore $a \sim a$ because 1 is a unit in R. Thus $\sim$ is reflexive.

Symmetry. We have $a \sim b \Rightarrow a=ub$ for some unit u in $R \Rightarrow u^{-1}a=u^{-1}ub \Rightarrow u^{-1}a=1b \Rightarrow u^{-1}a=b \Rightarrow b=u^{-1}a \Rightarrow b \sim a$ because u^{-1} is also a unit in R. Therefore $\sim$ is symmetric.

Transitivity. Let $a \sim b$, $b \sim c$. Then $a=ub$ and $b=vc$ for some units $u, v \in R$. This gives $a=u(vc)=(uv)c$. But the product of two units in R is again a unit in R. Therefore uv is a unit in R and so $a=(uv)c \Rightarrow a \sim c$. Therefore $\sim$ is transitive.

Hence $\sim$ is an equivalence relation in R.

Theorem 3. *Let D be an integral domain with unity element 1. Two non-zero elements $a, b \in D$ are associates if and only if $a \mid b$ and $b \mid a$.*

Proof. Let two non-zero elements a, b be associates of each other in D. Then $a=bu$ where u is a unit in R.

Now $a=bu \Rightarrow b \mid a$.

Again $a=bu \Rightarrow au^{-1}=buu^{-1} \Rightarrow au^{-1}=b \Rightarrow b=au^{-1} \Rightarrow a \mid b$.

Thus a and b are associates $\Rightarrow a \mid b$ and $b \mid a$.

Conversely, let $a \mid b$ and $b \mid a$. Then to prove that a and b are associates.

We have $a \mid b \Rightarrow \exists\, c \in D$ such that $b=ac$.

Similarly $b \mid a \Rightarrow \exists\, d \in D$ such that $a=bd$.

$\therefore \quad b=ac=(bd)\,c=b\,(dc)$.

$\therefore \quad b1=b\,(dc)$. $\qquad [\because\ b1=b]$

or $\quad b\,(1-dc)=0$

or $\quad (1-dc)=0 \qquad [\because\ b \neq 0$ and D is without zero divisors]

or $\quad 1=dc$.

$\therefore \quad$ Both c and d are units in D.

Thus $a=bd$ where d is a unit in D. Hence a and b are associates.

Note. In a field any two non-zero elements are associates.

Proper and Improper Divisors. Definition. *Let D be an integral domain with unity element 1. Let a be any non-zero element of D. Then the units of D and associates of a are always divisors of a. These are called improper or trivial divisors of a. Any other divisors of a are called proper or non-trivial divisors of a.*

In the integral domain of integers ± 1, ± 6 are trivial divisors of 6. But ± 2, ± 3 are proper divisors of 6.

Prime Elements. Definition. (Raj. 77). *Let D be an integral domain with unity element* 1. *A non-zero non-unit element $a \in D$, having only trivial divisors, is called a* **prime** *or* **irreducible** *element of D.* An element $0 \neq b \in D$ having proper divisors is called a **reducible** or **composite** element of D. From this definition it is obvious that if p is a prime element of D and if $p=xy$, where $x, y \in D$, then one of x or y must be a unit in D.

Also $0 \neq b \in D$ is a composite element of D if and only if we can find two elements $x, y \in D$ such that $b=xy$ and none of x and y is a unit in D.

Greatest Common Divisor. Definition. *Let R be a commutative ring. If $a, b \in R$ then $0 \neq d \in R$ is said to be a greatest common divisor of a and b if*

(*i*) *$d \mid a$ and $d \mid b$.*

(*ii*) *Whenever $c \mid a$ and $c \mid b$ then $c \mid d$.* **(Madras 1983)**

We shall use the notation $d=(a, b)$ to denote that d is a greatest common divisor of a and b.

Now suppose $a, b \in D$ where D is an integral domain with unity element 1. Let a, b possess a greatest common divisor.

If d_1, d_2 are two greatest common divisors of a and b, we have

$$d_1 \mid d_2 \text{ and } d_2 \mid d_1$$

$$\Rightarrow d_1 \text{ and } d_2 \text{ are associates.}$$

Thus in an integral domain with unity in case a greatest common divisor of a and b exists, it is unique apart from the distinction between associates.

Relatively Prime Elements. Definition. (Raj. 1967). *Let D be an integral domain with unity element* 1. *Two elements $a, b \in D$ are said to be relatively prime if their greatest common divisor is a unit of D.*

But any associate of a greatest common divisor is a greatest common divisor. Also the unity element 1 is an associate of any unit. Therefore if a, b are relatively prime we may assume that a greatest common divisor of a and b is 1 *i.e.*, $(a, b)=1$.

§ 20. **Polynomial Rings.** While studying algebra in high school classes we are introduced polynomials. We know that expressions of the type $3x^2-4x+5$, $x+7$, $9x^3-\frac{3}{2}x^2+4x+5$ etc. are called polynomials in the *indeterminate* x. In place of x we can use other letters like y, z etc. to denote these polynomials. Now we shall define polynomials over an arbitrary ring.

Definition. *Let R be an arbitrary ring and let x, called an* **indeterminate,** *be any symbol not an element of R By a polynomial in x over R is meant an expression of the form*

$$f(x) = a_0x^0 + a_1x + a_2x^2 + \ldots\ldots, \text{ where}$$

$a_0, a_1, a_2, \ldots$ are elements of R and only a finite number of them are not equal to 0, the zero element of R. **(Agra 1970; Meerut 73)**

Here x is an indeterminate. We could have used any other letter, say, y in place of x. Also a_0x^0, a_1x, a_2x^2 etc. are called terms of the polynomial and a_0, a_1, a_2 etc. are called coefficients of these terms. All these coefficients are elements of R. The number of terms in the polynomial $f(x)$ will be infinite but except a finite number of terms, the coefficients of all the remaining terms will be equal to 0, the zero element of the ring. The symbol '+' connecting various terms in $f(x)$ has no connection with the addition of the ring R. This symbol has been used here only to connect different terms. Also x is not an element of R. The powers of x are nothing to do with the powers of an element of R. The different powers of x only tell us the ordered place of different coefficients. There is no harm if we represent this polynomial $f(x)$ by the infinite ordered set $(a_0, a_1, a_2, \ldots\ldots)$ where $a_0, a_1, a_2, \ldots\ldots$ are elements of R and only a finite number of them are not equal to zero. Since from high school classes we represent a polynomial with an indeterminate x, therefore we have preferred this way to represent polynomials.

Set of all polynomials over a ring. *Let R be an arbitrary ring and x an indeterminate. The set of all polynomials $f(x)$,*

$$f(x) = \sum_{n=0}^{\infty} a_nx^n = a_0x^0 + a_1x + a_2x^2 + \cdots$$

where the a's are elements of the ring R and only a finite number of them are not equal to zero, is called $R\,[x]$.

We shall make a ring out of $R\,[x]$. Then $R\,[x]$ will be called the ring of all polynomials over the ring R. For this we shall define **equality, addition,** and **multiplication** of two elements of $R\,[x]$.

Definition. *Suppose R is an arbitrary ring and*

$$f(x) = a_0x^0 + a_1x + a_2x^2 + a_3x^3 + \ldots$$

and $g(x)=b_0x^0+b_1x+b_2x^2+b_3x^3+\ldots$
are any elements of $R[x]$. *Then*

(*a*) $f(x)=g(x)$ *if and only if* $a_n=b_n$ $\forall$ *non-negative integer* n. *Thus two polynomials are equal iff their corresponding coefficients are equal.*

(*b*) $f(x)+g(x)=c_0x^0+c_1x+c_2x^2+c_3x^3+\ldots$ *where* $c_n=a_n+b_n$ *for every non-negative integer* n *Thus in order to add two polynomials we should add the coefficients of like powers of* x.

Since $c_n \in R$ and only a finite number of c's can be not equal to zero, therefore $f(x)+g(x)$ is an element of $R[x]$. Thus $R[x]$ is closed with respect to addition of polynomials as defined above.

(*c*) $f(x)\,g(x)=d_0x^0+d_1x+d_2x^2+d_3x^3+\ldots$
where $d_n=a_0b_n+a_1b_{n-1}+a_2b_{n-2}+\ldots+a_nb_0$
for every non-negative integer n. We can write $d_n=\sum_{i+j=n} a_ib_j$

where by this summation we mean the sum of all the products of the type a_ib_j with i and j non-negative integers whose sum is n.

Since $d_n \in R$ and only a finite number of d's can be not equal to zero, therefore $f(x)\,g(x)$ is an element of $R[x]$. Thus $R[x]$ is closed with respect to multiplication of polynomials as defined above.

We have $d_0=a_0b_0$, $d_1=a_0b_1+a_1b_0$,
$d_2=a_0b_2+a_1b_1+a_2b_0$, $d_3=a_0b_3+a_1b_2+a_2b_1+a_3b_0$ and so on.

Therefore in order to multiply two polynomials $f(x)$ and $g(x)$, we should first write

$f(x)\,g(x)=(a_0x^0+a_1x+a_2x^2+a_3x^3+\ldots)\,(b_0x^0+b_1x+b_2x^2+\ldots)$.

Now we should multiply different powers of the indeterminate x and using the relation $x^ix^j=x^{i+j}$ we should collect coefficients of different powers of x.

Zero Polynomial. The polynomial

$$f(x)=\Sigma a_nx^n=a_0x^0+a_1x+a_2x^2+a_3x^3+\ldots$$

in which all the coefficients $a_0, a_1, a_2, \ldots$ are equal to 0 is called the zero polynomial over the ring R.

Degree of a Polynomial. **(Meerut 1978)**

Let $f(x)=a_0x^0+a_1x+a_2x^2+a_3x^3+\ldots+a_nx^n+\ldots$
be a polynomial over an arbitrary ring R. *We say that* n *is the degree of the polynomial* $f(x)$ *if and only if* $a_n\neq 0$ *and* $a_m=0$ *for all*

$m>n$. We shall write $\deg f(x)$ to denote the degree of $f(x)$. Thus the degree of $f(x)$ is the largest non-negative integer i for which the ith coefficient of $f(x)$ is not 0. If in the polynomial $f(x)$, a_0 (*i.e.* the coefficient of x^0) is not 0 and all the other coefficients are 0, then according to our definition, the degree of $f(x)$ will be zero. Also according to our definition, if there is no non-zero coefficient in $f(x)$, then its degree will remain undefined. Thus we do not define the degree of the zero polynomial. Also it is obvious that every non-zero polynomial will possess a unique degree.

Note. If $f(x)=a_0x^0+a_1x+a_2x^2+\ldots+a_nx^n+\ldots$ is a polynomial of degree n *i.e.*, if $a_n\neq 0$ and $a_m=0$ for all $m>n$, then it is convenient to write $f(x)=\sum_{i=0}^{n} a_ix^i=a_0x^0+a_1x+a_2x^2+\ldots+a_nx^n$. It will remain understood that all the terms in $f(x)$ which follow the term a_nx^n, have zero coefficients. Also we shall call a_nx^n as the **leading term** and a_n as the **leading coefficient** of the polynomial. The term a_0x^0 is called the **constant term** and a_0 is called the zeroth coefficient of $f(x)$. For example $f(x)=2x^0+3x-4x^2+5x^3-8x^4$ is a polynomial of degree 4 over the ring of integers. Here -8 is the leading coefficient and 2 is the zeroth coefficient. The coefficients of all terms which contain powers of x greater than 4 will be regarded as zero. Similarly $g(x)=3x^0$ is a polynomial of degree zero over the ring of integers. In this polynomial the coefficients of x, x^2, x^3, .. are all equal to zero. The zero polynomial over an arbitrary ring R will be represented by $0x^0$.

Set of constant polynomials over a ring. *Let R be an arbitrary ring and $R[x]$ the set of all polynomials over R Let R' denote the set of all polynomials over R whose coefficients are all zero except for the constant term, which may be either zero or non-zero. That is,*

$$R'=\{ax^0 : a\in R\}.$$

Then R' will be called as the set of constant polynomials in $R[x]$.

Thus all the polynomials of degree 0 as well as the zero polynomial will be called as constant polynomials.

Example 1. *Add and multiply the following polynomials over the ring of integers :*

$$f(x)=2x^0+5x+3x^2-4x^3,\ g(x)=3x^0+4x-x^3+5x^4.$$

Solution. By our definition of the sum of two polynomials, we have

$$f(x)+g(x)=(2+3)\ x^0+(5+4)x+(3+0)\ x^2$$
$$+(-4-1)\ x^3+(0+5)\ x^4$$
$$=5x^0+9x+3x^2-5x^3+5x^4.$$

Also $f(x)\ g(x)=(2x^0+5x+3x^2-4x^3)\ (3x^0+4x-x^3+5x^4)$

$$=6x^0+(8+15)\ x+(20+9)\ x^2+(-2+12-12)\ x^3$$
$$+(10-5-16)\ x^4+(25-3)\ x^5+(15+4)\ x^6-20x^7$$
$$=6x^0+23x+29x^2-2x^3-11x^4+22x^5+19x^6-20x^7.$$

Example 2. *Add and multiply the following polynomials over the ring* $(I_6,\ +_6,\ \times_6)$:

$$f(x)=2x^0+5x+3x^2,\ g(x)=1x^0+4x+2x^3.$$

Solution. $f(x)+g(x)=(2+_6 1)\ x^0+(5+_6 4)\ x$

$$+(3+_6 0)\ x^2+(0+_6 2)\ x^3$$
$$=3x^0+3x+3x^2+2x^3.$$

Also $f(x)\ g(x)=(2x^0+5x+3x^2)\ (1x^0+4x+2x^3)$

$$=(2\times_6 1)\ x^0+[(2\times_6 4)+_6(5\times_6 1)]x+[(5\times_6 4)+_6 (3\times_6 1)]\ x^2$$
$$+[(2\times_6 2)+_6 (3\times_6 1)]\ x^3+(5\times_6 2)\ x^4+(3\times_6 2)x^5$$
$$=2x^0+(2+_6 5)\ x+(2+_6 3)\ x^2+(4+_6 0)\ x^3+4x^4+0x^5$$
$$=2x^0+1x+5x^2+4x^3+4x^4.$$

Note. Here degree of $f(x)=2$, degree of $g(x)=3$ and degree of $f(x)\ g(x)=4$. The point to note is that degree of $f(x)\ g(x)$ may be less than the sum of the degrees of $f(x)$ and $g(x)$.

§ 21. Degree of the sum and the product of two polynomials.

Theorem. *Let $f(x)$ and $g(x)$ be two non-zero polynomials over an arbitrary ring R. Then*

(*i*) *deg* $[f(x)+g(x)] \leqslant$ *Max* [*deg* $f(x)$. *deg* $g(x)$],

if $f(x)+g(x)\neq 0$.

(*ii*) *deg* $[f(x)\ g(x)] \leqslant$ *deg* $f(x)$ + *deg* $g(x)$ *if* $f(x)\ g(x)\neq 0$.

(Meerut 76, 78)

Proof. Let $f(x)=a_0x^0+a_1x+a_2x^2+\ldots+a_nx^n,\ a_n\neq 0$ and $g(x)=b_0x^0+b_1x+b_2x^2+\ldots+b_mx^m,\ b_m\neq 0$ be two elements of $R\,[x]$.

Here deg $f(x)=n$ and deg $g(x)=m$.

From our definition of the sum of two polynomials, it is obvious that if $f(x)+g(x)\neq 0$, then

$$\deg[f(x)+g(x)]=\begin{cases} \max(n,\ m) \text{ if } n\neq m \\ n \text{ if } n=m \text{ and } a_n+b_m\neq 0 \\ <n \text{ if } n=m \text{ and } a_n+b_m=0. \end{cases}$$

Again $f(x)\ g(x)=(a_0b_0)\ x^0+(a_0b_1+a_1b_0)\ x+\ldots+a_nb_mx^{n+m}$.

Suppose $f(x)\ g(x)\neq 0$. Then $f(x)\ g(x)$ has a unique degree.

If $a_n b_m \neq 0$, then deg $[f(x)\, g(x)] = n+m = \deg f(x) + \deg g(x)$.

Also if $a_n b_m = 0$, then deg $[f(x)\, g(x)] < n+m$.

Cor. 1. Important. *Suppose D is an integral domain and $f(x)$, $g(x)$ are two non-zero elements of $D[x]$. Then*

$$\deg [f(x)\, g(x)] = \deg f(x) + \deg g(x).$$

Proof. Since $a_n \neq 0$, $b_m \neq 0$, therefore $a_n b_m \neq 0$ because in an integral domain the product of two non-zero elements cannot be zero. Hence deg $[f(x)\, g(x)] = m+n$.

Cor. 2. *If F is a field and $f(x)$, $g(x)$ are two non-zero elements of $F[x]$, then deg $[f(x)\, g(x)] = \deg f(x) + \deg g(x)$.*

(Meerut 1973; Kanpur 80)

Proof. Since a field is also free from zero divisors, therefore $a_n b_m \neq 0$ when $a_n \neq 0$ and $b_m \neq 0$ Hence the result.

§ 22. Ring of Polynomials.

Theorem. *The set $R[x]$ of all polynomials over an arbitrary ring R is a ring with respect to addition and multiplication of polynomials.* **(Meerut 1973, 80; Sagar 77)**

Proof. Let $f(x), g(x) \in R[x]$. Then $f(x)+g(x)$ and $f(x)\, g(x)$ are also polynomials over R. Therefore $R[x]$ is closed with respect to addition and multiplication of polynomials.

Now let

$f(x) = \Sigma a_i x^i = a_0 x^0 + a_1 x + a_2 x^2 + \ldots$, $g(x) = b_0 x^0 + b_1 x + b_2 x^2 + \ldots$, $h(x) = c_0 x^0 + c_1 x + c_2 x^2 + \ldots$ be any arbitrary elements of $R[x]$.

Commutativity of addition. We have

$$f(x) + g(x) = (a_0 + b_0)\, x^0 + (a_1 + b_1)\, x + (a_2 + b_2)\, x^2 + \ldots$$
$$= (b_0 + a_0)\, x^0 + (b_1 + a_1)\, x + (b_2 + a_2)\, x^2 + \ldots = g(x) + f(x).$$

Associativity of addition. We have

$$[f(x) + g(x)] + h(x) = \Sigma\, (a_i + b_i)\, x^i + \Sigma\, c_i x^i = \Sigma\, [(a_i + b_i) + c_i]\, x^i$$
$$= \Sigma\, [a_i + (b_i + c_i)]\, x^i = \Sigma\, a_i x^i + \Sigma\, (b_i + c_i)\, x^i = f(x) + [g(x) + h(x)].$$

Existence of additive indentity. Let $0\,(x)$ be the zero polynomial over R *i.e.*, $0(x) = 0x^0 + 0x + 0x^2 + \ldots$

$$\text{Then } f(x) + 0\,(x) = (a_0 + 0)\, x^0 + (a_1 + 0)\, x + (a_2 + 0)\, x^2 + \ldots$$
$$= a_0 x^0 + a_1 x + a_2 x^2 + \ldots = f(x).$$

$\therefore$ the zero polynomial $0(x)$ is the additive identity.

Existence of additive inverse. Let $-f(x)$ be the polynomial over R defined as $-f(x) = (-a_0)\, x^0 + (-a_1)\, x + (-a_2)\, x^2 + \ldots$

Then $-f(x) + f(x) = (-a_0 + a_0)\, x^0 + (-a_1 + a_1)\, x + (-a_2 + a_2)\, x^2 + \ldots$
$= 0x^0 + 0x + 0x^2 + \ldots = 0\,(x) =$ the additive identity.

$\therefore$ each member of $R[x]$ possesses additive inverse.

Associativity of Multiplication. We have

$$f(x)\,g(x)=(a_0x^0+a_1x+a_2x^2+\ldots)\,(b_0x^0+b_1x+b_2x^2+\ldots)$$
$$=d_0x^0+d_1x+d_2x^2+\ldots+d_lx^l+\ldots,\text{ where } d_l=\sum_{i+j=l} a_ib_j.$$

Now $[f(x)\,g(x)]\,h(x)$

$$=(d_0x^0+d_1x+d_2x^2+\ldots)\,(c_0x^0+c_1x+c_2x^2+\ldots)$$
$$=e_0x^0+e_1x+e_2x^2+\ldots+e_nx^n+\ldots,$$

where $e_n=$ the coeff. of x^n in $[f(x)\,g(x)]\,h(x)$

$$=\sum_{l+k=n} d_lc_k=\sum_{l+k=n}\Big[\Big(\sum_{i+j=l} a_ib_j\Big)c_k\Big]=\sum_{i+j+k=n} a_ib_jc_k.$$

Similarly we can show that the coeff. of x^n in

$$f(x)\,[g(x)\,h(x)]=\sum_{i+j+k=n} a_ib_jc_k.$$

Thus $[f(x)\,g(x)]\,h(x)=f(x)\,[g(x)\,h(x)]$ since corresponding coefficients in these two polynomials are equal.

Distributivity of Multiplication with respect to addition. We have $f(x)\,[g(x)+h(x)]$

$$=(a_0x^0+a_1x+a_2x^2+\ldots)\,[(b_0+c_0)\,x^0+(b_1+c_1)\,x+(b_2+c_2)\,x^2+\ldots].$$

If n is any non-negative integer, then the coefficient of x^n in $f(x)\,[g(x)+h(x)]$

$$=\sum_{i+j=n} a_i\,(b_j+c_j)=\sum_{i+j=n}(a_ib_j+a_ic_j)=\sum_{i+j=n} a_ib_j+\sum_{i+j=n} a_ic_j$$

$=$ Coeff of x^n in $f(x)\,g(x)+$ coeff. of x^n in $f(x)\,h(x)$

$=$ Coeff. of x^n in $[f(x)\,g(x)+f(x)\,h(x)]$.

$\therefore\ f(x)\,[g(x)+h(x)]=f(x)\,g(x)+f(x)\,h(x).$

Similarly we can prove the right distributive law.

Hence $R[x]$ is a ring. This is called the ring of all polynomials over R. The zero element of this ring is the zero polynomial

$$0x^0+0x+0x^2+0x^3+\ldots$$

§ 23. R as a subset of $R[x]$ or Imbedding of R into $R[x]$.

Theorem. *If R is an arbitrary ring and R' is the set of constant polynomials in $R[x]$, then R' is isomorphic to R.* **(I.A.S. 1974)**

Proof. We have $R'=\{ax^0+0x+0x^2+0x^3+\ldots$ such that $a\in R\}$.

Let $\phi : R\to R'$ such that

$$\phi(a)=ax^0+0x+0x^2+0x^3+\ldots\ \forall\ a\in R.$$

ϕ is one-one since

$\phi(a)=\phi(b)\Rightarrow ax^0+0x+0x^2+\ldots=bx^0+0x+0x^2+\ldots\Rightarrow a=b.$

Also ϕ is obviously onto R'.

Again $\phi(a+b)=(a+b)\,x^0+0x+0x^2+0x^3+\ldots$

$=[ax^0+0x+0x^2+\ldots]+[bx^0+0x+0x^2+\ldots]=\phi\,(a)+\phi\,(b).$

Also $\phi(ab)=abx^0+0x+0x^2+0x^3+\ldots$
$=(ax^0+0x+0x^2+\ldots)(bx^0+0x+0x^2+\ldots)=\phi(a)\,\phi(b)$.

$\therefore$ ϕ is an isomorphism of R onto R'. Hence $R\cong R'$.

Since R is isomorphic to R', therefore in $R[x]$ we can identify R' by R *i.e.* all the constant polynomials in $R[x]$ can be replaced by the corresponding elements of R. This replacement will not affect the addition and multiplication of polynomials.

Hence in future we shall write 0 in place of the zero polynomial. If $ax^0+0x+0x^2+\ldots$ is any constant polynomial in $R[x]$, then we shall simply write a in place of this polynomial.

Also in place of $a_0x^0+a_1x+a_2x^2+\ldots\ldots$ we shall write $a_0+a_1x+a_2x^2+\ldots\ldots$ If we are to multiply $f(x)$ by a constant polynomial $ax^0+0x+0x^2+\ldots$, then we shall write $af(x)$ in place of $(ax^0+0x+\ldots)\,f(x)$.

§ 24. Polynomials over an integral domain.

Theorem. *If D is an integral domain, then the polynomial ring $D[x]$ is also an integral domain.*

(I.A.S. 1974 ; Rajasthan 74 ; Jabalpur 70 ; Allahabad 80 ; Gujrat 71 ; Marathwada 72 ; Meerut 78, 84 P)

Proof. Let D be a commutative ring without zero divisors and with unity element 1. As proved in § 22, $D[x]$ is also a ring. To prove that $D[x]$ is an integral domain, we should prove that (i) $D[x]$ is commutative, (ii) is without zero divisors and (iii) possesses the unity element.

$D[x]$ is commutative. Let $f(x)=a_0+a_1x+a_2x^2+\ldots\ldots$ and $g(x)=b_0+b_1x+b_2x^2+\ldots$ be any two elements of $D[x]$.

If n is any non-negative integer, then the coefficient of x^n in $f(x)\,g(x)$ is $=\sum_{i+j=n} a_ib_j=\sum_{i+j=n} b_ja_i$, since D is commutative

$=$ Coefficient of x^n in $g(x)f(x)$.

$\therefore$ $f(x)g(x)=g(x)\,f(x)$. Hence $D[x]$ is a commutative ring.

If 1 is the unity element of D, then the constant polynomial $1+0x+0x^2+0x^3+\ldots$ is the unity element of $D[x]$. We have
$[a_0+a_1x+a_2x^2+\ldots][1+0x+0x^2+0x^3+\ldots]$
$=(a_01)+(a_11)\,x+(a_21)\,x^2+\ldots=a_0+a_1x+a_2x^2+\ldots$

$\therefore$ the polynomial $1+0x+0x^2+\ldots$ or simply 1 is the unity element of $D[x]$.

$D[x]$ is without zero divisors. Let

$$f(x)=a_0+a_1x+a_2x^2+\dots+a_mx^m,\ a_m\neq 0$$
$$g(x)=b_0+b_1x+b_2x^2+\dots+b_nx^n,\ b_n\neq 0$$

be two non-zero elements of $D[x]$.

Then $f(x)\ g(x)$ cannot be a zero polynomial *i.e.*, the zero element of $D\ [x]$. The reason is that at least one coefficient of $f(x)\ g(x)$ namely a_mb_n of x^{m+n} is $\neq 0$ because a_m, b_n are non-zero elements of D and D is without zero divisors.

Hence $D[x]$ is an integral domain.

Theorem 2. *If R is an integral domain with unity element, then any unit in R[x] must already be a unit in R.*

Proof. If R is an integral domain with unity element 1, then $R[x]$ is also an integral domain with unity element. Further the constant polynomial 1 is the unity element of $R[x]$. Let $f(x)$ be a unit in $R[x]$, *i.e.*, let $f(x)$ be an inversible element of $R[x]$. Let $g(x)$ be the inverse of $f(x)$ in $R[x]$. Then

$$f(x)\ g(x)=1$$

$\Rightarrow \deg\,[f(x)\ g(x)]=0$ [$\because$ degree of the constant polynomial 1 is 0]

$\Rightarrow \deg f(x)+\deg g(x)=0 \Rightarrow \deg f(x)=0,\ \deg g(x)=0$

$\Rightarrow$ both $f(x)$ and $g(x)$ are constant polynomials in $R[x]$.

Let $f(x)=a\in R$ and $g(x)=b\in R$. Then $ab=1 \Rightarrow a$ is a unit in R. Thus any unit in $R[x]$ must already be a unit in R.

Note. If $a\in R$ is a unit in R, then $a\in R[x]$ is also a unit in $R[x]$. If b is the inverse of a in R, then the constant polynomial b is the inverse of a in $R[x]$.

§ 25. Polynomials over a field.

Theorem 1. *If F is a field, then the set F[x] of all polynomials over F is an integral domain.*

(Aligarh 1966 ; Bombay 70 ; Allahabad 66 ; Meerut 69)

Proof. Every field is an integral domain. So give the same proof as we have given in § 22 and then in § 24.

We shall call the set $F[x]$ as the polynomial domain over the field F.

Theorem 2. *The polynomial domain F[x] over a field F is not a field.*

Proof. In order to show that $F[x]$ is not a field, we should show that there exists a non-zero element of $F[x]$ which has no multiplicative inverse. Let $f(x)$ be any member of $F[x]$ such that $\deg f(x)$ is greater than zero. The inverse of $f(x)$ cannot be the zero polynomial because the product of $f(x)$ and the zero polynomial will be equal to the zero polynomial and not equal to the unity element of $F[x]$ which is the polynomial $1+0x+0x^2+\dots$. Suppose now $g(x)$ is any non-zero polynomial. Then F being a field, we have

$$\deg [f(x)\, g(x)] = \deg f(x) + \deg g(x) > 0 \text{ because } \deg f(x) > 0 \text{ and } \deg g(x) \geqslant 0.$$

The degree of the unity element of $F[x]$ is 0. Hence $f(x)\, g(x)$ cannot be equal to the unity element of $F[x]$. Thus $f(x)$ does not possess multiplicative inverse.

$\therefore$ $F[x]$ is not a field.

Important. The only inversible elements of $F[x]$ are constant polynomials excluding the zero polynomial. No member of $F[x]$ whose degree is greater than 0 is inversible.

§ 26. Ring of polynomials in n variables over an integral domain.

Definition. *Let R be an integral domain. Then the ring of polynomials in the n-variables $x_1, \dots, x_n$ over R is denoted by $R\,[x_1, \dots, x_n]$ and is defined as follows :*

Let $R_1 = R\,[x_1]$, *the polynomial ring in x_1 over R,*

$R_2 = R_1\,[x_2]$, *the polynomial ring in x_2 over R_1,*

$R_3 = R_2\,[x_3]$, *the polynomial ring in x_3 over R_2,*

$R_n = R_{n-1}\,[x_n]$, *the polynomial ring in x_n over R_{n-1}.*

Then R_n is called the ring of polynomials in $x_1, \dots, x_n$ over R and we write $R_n = R\,[x_1, \dots, x_n]$.

Theorem 1. *If R is an integral domain, then so is $R\,[x_1, \dots, x_n]$.*

Proof. If R is an integral domain then $R_1 = R\,[x_1]$ is also an integral domain. Now R_1 is an integral domain implies that $R_2 = R_1\,[x_2] = R\,[x_1, x_2]$ is also an integral domain. Continuing this process a finite number of times we see that $R\,[x_1, \dots, x_n]$ is an integral domain.

Theorem 2. *If F is a field, then $F\,[x_1, \dots, x_n]$ is an integral domain.*

Proof. If F is a field, then $F_1 = F\,[x_1]$ is an integral domain.

Now F_1 is an integral domain implies that $F_2 = F_1\,[x_2] = F\,[x_1, x_2]$ is also an integral domain. Continuing this process a finite number of times we see that $F\,[x_1, \dots, x_n]$ is an integral domain.

Note. If F is a field, then $F\,[x_1, \dots, x_n]$ is an integral domain.

Now we can construct the field of quotients of the integral domain $F[x_1, ..., x_n]$. This field is called *the field of rational functions in* $x_1, ..., x_n$ *over* F and is denoted by $F(x_1, ..., x_n)$.

§ 27. Divisibility of Polynomials over a field.

Suppose F is a field. Then $F[x]$ is an integral domain. If $a(x) \neq 0$ and $f(x)$ are elements of $F[x]$, then $a(x)$ is a *divisor* (*or factor*) of $f(x)$ if and only if there is a polynomial $b(x)$ in $F[x]$ such that $f(x)=a(x)\, b(x)$. Symbolically we write $a(x) \mid f(x)$.

A **unit** is an element of $F[x]$ which has multiplicative inverse. All the polynomials of zero degree belonging to $F[x]$ are units of $F[x]$. Thus the non-zero elements of F are the only units of $F[x]$.

If $f(x)$ and $g(x)$ are polynomials in $F[x]$, then we call $f(x)$ and $g(x)$ *associates* if $f(x)=c\, g(x)$ for some $0 \neq c \in F$. It can be easily proved that two non-zero polynomials $f(x)$ and $g(x)$ in $F[x]$ are associates if and only if $f(x) \mid g(x)$ and $g(x) \mid f(x)$.

If $f(x)$ is any non-zero polynomial in $F[x]$, then $f(x)$ is always divisible by its associates and by all units of $F[x]$. These divisors of $f(x)$ are called its *improper divisors*. All other divisors of $f(x)$, if there are any, are called its *proper divisors*.

Definition of an irreducible polynomial over a field. *Let F be a field and $f(x)$ be a non-zero and non-unit polynomial in $F[x]$ i.e., $f(x)$ be a polynomial of positive degree. Then $f(x)$ is said to be irreducible over F (or prime) if it has no proper divisors in $F[x]$; $f(x)$ is reducible over F if it has a proper divisor in $F[x]$.*

(Meerut 1983P)

Thus a positive degree polynomial $f(x)$ in $F[x]$ is irreducible over F if whenever $f(x)=a(x)\, b(x)$ with $a(x), b(x) \in F[x]$ then one of $a(x)$ or $b(x)$ is a unit in $F[x]$ *i.e.*, has degree 0. Also $f(x)$ is reducible over F if and only if we can find two polynomials $a(x)$ and $b(x)$ in $F[x]$ such that $f(x)=a(x)\, b(x)$ and none of $a(x)$ and $b(x)$ is a unit in $F[x]$ *i.e.*, has degree 0.

Irreducibility depends on the field. The polynomial x^2-2 is irreducible over the field of rational numbers while it is reducible over the field of real numbers, since $x^2-2=(x+\sqrt{2})(x-\sqrt{2})$.

The polynomial x^2+1 is irreducible over the field of real numbers while it is reducible over the field of complex numbers since $x^2+1=(x+i)(x-i)$.

Monic polynomials. Definition.

Let $f(x)=a_0+a_1x+...+a_nx^n$, with $a_n \neq 0$, be a polynomial in

$F[x]$ over an arbitrary field F. If the leading coefficient a_n of $f(x)$ is equal to 1, the unity element of F then the polynomial $f(x)$ will be called monic.

The polynomial $2x-3x^3$ over the field of real numbers is not monic since its leading coefficient is -3. But the polynomial x^3-3x+4 over the field of real numbers is monic since its leading coefficient is 1.

Greatest common divisor of two polynomials over a field

Definition. *Suppose F is any field. Let $f(x)$ and $g(x)$ be two elements of $F[x]$. A greatest common divisor of $f(x)$ and $g(x)$ is a non-zero polynomial $d(x)$ such that*

(i) $d(x) \mid f(x)$ and $d(x) \mid g(x)$.

(ii) If $c(x)$ is a polynomial such that $c(x) \mid f(x)$ and $c(x) \mid g(x)$, then $c(x) \mid d(x)$.

Relatively prime polynomials *Two polynomials $f(x)$ and $g(x) \in F[x]$ are said to be relatively prime if their greatest common divisor is 1, the unity element of F.*

§ 28 Division Algorithm for polynomials over a field

Theorem. *Let $f(x)$, $g(x)\neq0$ be any two polynomials of the polynomial domain $F[x]$, over the field F. Then there exist uniquely two polynomials $q(x)$ and $r(x)$ in $F[x]$ such that*

$$f(x)=q(x)\,g(x)+r(x)$$

where either $r(x)=0$ or $\deg r(x) < \deg g(x)$.

(Allahabad 1970; Vikram 76; Kerala 70; Bombay 70; Meerut 74, 76)

Proof. Suppose

$$f(x)=a_0+a_1x+a_2x^2+\dots+a_mx^m,\ a_m\neq0$$

and $$g(x)=b_0+b_1x+b_2x^2+\dots+b_nx^n,\ b_n\neq0.$$

If degree m of $f(x)$ is smaller than the degree n of $g(x)$ or if $f(x)=0$, then we are nothing to prove Because we can always write $f(x)=0.\ g(x)+f(x)$. So in this case $q(x)=0$, $r(x)=f(x)$ and we have either $f(x)=0$ or $\deg f(x) < \deg g(x)$.

Now let us assume that $m > n$. In this case we shall prove the theorem by induction on m *i.e.*, degree of $f(x)$.

If $m=0$, then $m \geq n \Rightarrow n=0$. Therefore $f(x)$ and $g(x)$ are both non-zero constant polynomials, $f(x)=a_0$, $a_0\neq0$, and $g(x)=b_0$, $b_0\neq0$. We have in this case

$$f(x)=a_0=(a_0b_0^{-1})\,b_0+0=(a_0b_0^{-1})\,g(x)+0.$$

Thus the theorem is true when $m=0$ or when the degree of $f(x)$ is less than 1.

We shall now assume that the theorem is true when $f(x)$ is a polynomial of degree less than m and then we shall show that it is also true if $f(x)$ is of degree m and then the proof will be complete by induction.

Let $f_1(x)=f(x)-(a_m b_n^{-1})\, x^{m-n}\, g(x)$ $\quad$ ..(1)

Obviously $\deg f_1(x) < m$. Therefore by our assumed hypothesis, there exist polynomials $s(x)$ and $r(x)$ such that

$$f_1(x)=s(x)\, g(x)+r(x),$$

where $\qquad r(x)=0$ or $\deg r(x) < \deg g(x)$.

Now putting the value of $f_1(x)$ in (1), we get

$$s(x)\, g(x)+r(x)=f(x)-(a_m b_n^{-1})\, x^{m-n}\, g(x)$$

or $$f(x)=[(a_m b_n^{-1})x^{m-n}+s(x)]\, g(x)+r(x).$$

If we write $q(x)$ in place of $(a_m b_n^{-1})\, x^{m-n}+s(x)$, we get

$$f(x)=q(x)\, g(x)+r(x)$$

where $\qquad s(x)=0$ or $\deg r(x) < \deg g(x)$.

This proves the existence of polynomials $q(x)$ and $r(x)$. Now to show that $q(x)$ and $r(x)$ are unique. Let us assume that

$$f(x)=q_1(x)\, g(x)+r_1(x)=q_2(x)\, g(x)+r_2(x).$$

Then $\qquad q_1(x)\, g(x)+r_1(x)=q_2(x)\, g(x)+r_2(x)$

or $\qquad [q_1(x)-q_2(x)]\, g(x)=r_2(x)-r_1(x).$ $\quad$ ...(2)

If $[q_1(x)-q_2(x)]\neq 0$, then $[q_1(x)-q_2(x)]\, g(x)$ cannot be equal to the zero polynomial because $g(x)\neq 0$ and $F[x]$ is without zero divisors. Also then the degree of $[q_1(x)-q_2(x)]\, g(x)$ is at least n, the degree of $g(x)$. But $r_2(x)-r_1(x)$ is either equal to the zero polynomial or else its degree is less than n because the degrees of $r_2(x)$ and $r_1(x)$ are both less than n. Hence the equality (2) among two polynomials holds only if

$$q_1(x)-q_2(x)=0 \text{ and } r_2(x)-r_1(x)=0$$

i.e., only if $q_1(x)=q_2(x)$ and $r_1(x)=r_2(x)$

$\therefore$ the polynomials $q(x)$ and $r(x)$ are unique.

Definition. *In the division algorithm, the polynomial $q(x)$ is called the* **quotient** *on dividing $f(x)$ by $g(x)$ and the polynomial $r(x)$ is called the* **remainder.**

Theorem. *A polynomial domain $F[x]$ over a field F is a principal ideal ring* (**Vikram 1976; Lucknow 70; Madras 78; Kanpur 71; Kumayun 77; Meerut 79, 81**)

Proof. Obviously $F[x]$ is a commutative ring with unity and without zero divisors. Therefore $F[x]$ is a principal ideal ring if every ideal in $F[x]$ is a principal ideal.

Let S be an arbitrary ideal of $F[x]$. If S is the null ideal, then $S=(0)$ *i.e.*, the ideal of $F[x]$ generated by 0. Therefore S is a principal ideal. So let us suppose that S is not a null ideal. Then there exist non-zero polynomials $f(x)$ in S. Let $g(x)$ be a polynomial of lowest degree m belonging to S. We shall show that S is the principal ideal generated by $g(x)$.

Let $f(x)$ be any arbitrary member of S. By division algorithm there exist two polynomials $q(x) \in F[x]$, $r(x) \in F[x]$, such that $f(x)=q(x)\, g(x)+r(x)$, where $r(x)=0$ or $\deg r(x) < \deg g(x)$.

Since S is an ideal, therefore

$$q(x) \in F[x],\ g(x) \in S \Rightarrow q(x)\, g(x) \in S.$$

Also $f(x) \in S,\ q(x)\, g(x) \in S \Rightarrow f(x)-q(x)\, g(x) \in S$.

But $f(x)-q(x)\, g(x)=r(x)$. Therefore $r(x) \in S$.

Now either $r(x)=0$ or $\deg r(x) < \deg g(x)$. But we have assumed that $g(x)$ is a polynomial of lowest degree belonging to S. Hence $\deg r(x)$ cannot be less than $\deg g(x)$. Therefore we must have $r(x)=0$. Then $f(x)=q(x)\, g(x)$. Thus $g(x) \in S$ is such that $f(x) \in S \Rightarrow f(x)=q(x)\, g(x)$ for some $q(x) \in F[x]$. Therefore S is a principal ideal of $F[x]$ generated by $g(x)$. Hence $F[x]$ is a principal ideal ring.

Important. A polynomial ring over an arbitrary field is a principal ideal ring. But a polynomial ring over an arbitrary ring is not a principal ideal ring as is obvious from the following example.

Example. *Show that the polynomial ring* **I** $[x]$ *over the ring of integers is not a principal ideal ring.* **(Meerut 1976, Andhra 75)**

Solution. To prove this statement we shall show that the ideal $(2, x)$ of the ring **I** $[x]$ generated by two elements 2 and x of **I** $[x]$ is not a principal ideal. Let $(2, x)$ be a principal ideal in **I** $[x]$. Then there will exist a non-zero element $g(x) \in$ **I** $[x]$ such that $(2, x)=(g(x))$.

Since $2 \in (g(x))$ and $x \in (g(x))$ therefore there will exist elements $\phi(x)$ and $\psi(x)$ belonging to **I** $[x]$ such that

$$2=\phi(x)\, g(x), \qquad \text{...(1)}$$

$$x=\psi(x)\, g(x). \qquad \text{...(2)}$$

From (1), we get $2x=[\phi(x)\, g(x)]\, x$ and from (2), we get

$$2x=2\psi(x)\, g(x).$$

$\therefore$ $2\psi(x)\, g(x)=x\phi(x)\, g(x)$ [$\because$ **I** $[x]$ is a commutative ring].

$\therefore$ $2\psi(x)=x\phi(x)$ since $g(x)\neq 0$, and **I** $[x]$ is without zero divisors.

Now $2\psi(x) = x\,\phi(x)$ implies that the coefficients of $\phi(x)$ must all be even integers. Therefore $\phi(x) = 2h(x)$ where $h(x)$ is some polynomial in $\mathbf{I}[x]$. Putting this value of $\phi(x)$ in (1) we get

$$2 = 2h(x)\,g(x)$$

or

$$1 = h(x)\,g(x).$$

Now $1 = h(x)\,g(x) \Rightarrow 1 \in (g(x))$. Therefore each element of $\mathbf{I}[x]$ will belong to $(g(x))$. Thus we have $\mathbf{I}[x] = (g(x)) = (2, x)$. Therefore each element of $\mathbf{I}[x]$ will belong to $(2, x)$. We shall show that $1 \notin (2, x)$ and this contradiction will mean that $(2, x)$ is not a principal ideal in $\mathbf{I}[x]$.

Now $1 \in (2, x) \Rightarrow$ we can write

$$1 = 2p(x) + xq(x)$$

where $p(x)$ and $q(x)$ are some elements in $\mathbf{I}[x]$.

Let $p(x) = a_0 + a_1x + a_2x^2 + \ldots$ and $q(x) = b_0 + b_1x + b_2x^2 + \ldots$

Then $1 = 2(a_0 + a_1x + a_2x^2 + \ldots) + x(b_0 + b_1x + b_2x^2 + \ldots)$

or $\quad 1 = 2a_0 + (2a_1 + b_0)\,x + (2a_2 + b_1)\,x^2 + \ldots$

This equality implies $1 = 2a_0$ where $a_0 \in \mathbf{I}$.

But for no integer a_0 we can have $1 = 2a_0$. Hence $1 \notin (2, x)$.

$\therefore$ $(2, x)$ is not a principal ideal in $\mathbf{I}[x]$.

§ 29. Euclidean Algorithm for polynomials over a field.

Theorem. *Let F be a field and $f(x)$ and $g(x)$ be any two polynomials in $F[x]$, not both of which are zero. Then $f(x)$ and $g(x)$ have a greatest common divisor $d(x)$ which can be expressed in the form*

$$d(x) = m(x)\,f(x) + n(x)\,g(x)$$

for polynomials $m(x)$ and $n(x)$ in $F[x]$. **(Meerut 1976)**

Proof Consider the set

$S = \{s(x)\,f(x) + t(x)\,g(x) : s(x), t(x) \in F[x]\}$. ...(1)

We claim that S is an ideal of $F[x]$. The proof is as follows :

Let $s_1(x)\,f(x) + t_1(x)\,g(x)$ and $s_2(x)\,f(x) + t_2(x)\,g(x)$ be any two elements of S.

Then $[s_1(x)\,f(x) + t_1(x)\,g(x)] - [s_2(x)\,f(x) + t_2(x)\,g(x)]$

$$= [s_1(x) - s_2(x)]\,f(x) + [t_1(x) - t_2(x)]\,g(x) \in S$$

since $s_1(x) - s_2(x)$ and $t_1(x) - t_2(x)$ are both members of $F[x]$.

Also if $\alpha(x)$ be any member of $F[x]$, then

$$\alpha(x)\,[s_1(x)\,f(x) + t_1(x)\,g(x)]$$

$$= [\alpha(x)\,s_1(x)]\,f(x) + [\alpha(x)\,t_1(x)]\,g(x) \in S.$$

Therefore S is an ideal of $F[x]$. Now every ideal in $F[x]$ is principal ideal. Therefore there exists an element $d(x)$ in S such that every element in S is a multiple of $d(x)$.

Since $d(x) \in S$, therefore from (1) we see that there exist elements $m(x)$, $n(x) \in F[x]$ such that

$$d(x)=m(x)\, f(x)+n(x)\, g(x).$$

Now $F[x]$ is a ring with unity element 1.

$\therefore$ Putting $s(x)=1$, $t(x)=0$ in (1), we see that $f(x) \in S$. Also putting $s(x)=0$, $t(x)=1$ in (1), we see that $g(x) \in S$.

Now $f(x)$, $g(x)$ are elements of S. Therefore they are both multiples of $d(x)$. Hence $d(x) \mid f(x)$ and $d(x) \mid g(x)$.

Now suppose $c(x) \mid f(x)$ and $c(x) \mid g(x)$.

Then $c(x) \mid [m(x) f(x)]$ and $c(x) \mid [n(x)\ g(x)]$. Therefore $c(x)$ is also a divisor of $m(x) f(x)+n(x)\, g(x)$ *i.e.*, $c(x)$ is a divisor of $d(x)$.

Thus $d(x)$ is a greatest common divisor of $f(x)$ and $g(x)$.

Note. If $d(x)$ is a greatest common divisor of $f(x)$ and $g(x)$ then any associate of $d(x)$ *i.e.*, $kd(x)$ where $0 \neq k \in F$ will also be a greatest common divisor of $f(x)$ and $g(x)$. In particular if $0 \neq b$ is the leading coefficient of the polynomial $d(x)$, then the monic polynomial $b^{-1}\, d(x)$ will also be a greatest common divisor of $f(x)$ and $g(x)$. *Often while defining greatest common divisor of two polynomials over a field we include one more condition in our definition that the greatest common divisor should be a monic polynomial.* The advantage of this extra condition is that now we shall get a unique greatest common divisor as shown below :

Suppose $d_1(x)$ and $d_2(x)$ are two monic polynomials and each is a greatest common divisor of $f(x)$ and $g(x)$. Then $d_1(x) \mid d_2(x)$ and $d_2(x) \mid d_1(x)$. Therefore $d_1(x)$ and $d_2(x)$ are associates and we have $d_1(x)=ud_2(x)$ for some $0 \neq u \in F$. Since $d_1(x)$ and $d_2(x)$ are both monic, therefore $u=1$.

§ 30. Unique Factorization Domain. Definition. *An integral domain, R, with unity element 1 is a unique factorization domain if*

(*a*) *any non-zero element in R is either a unit or can be written as the product of a finite number of irreducible (prime) elements of R ;*

(*b*) *the decomposition in part (a) is unique upto the order and associates of the irreducible elements.*

(Punjab 1968; Madurai 78; Banaras 64; Meerut 70)

Thus if R is a unique factorization domain and if $a \neq 0$ is a non-unit in R, then a can be expressed as a product of a finite number of prime elements of R. Also if

$$a = p_1 p_2 p_3 \ldots p_n = p_1' p_2' p_3' \ldots p_m'$$

where the p_i and p_j' are prime elements of R, then $m = n$ and each p_i, $1 \leqslant i \leqslant n$ is an associate of some p_j', $1 \leqslant j \leqslant m$ and conversely each p'_k is an associate of some p_i.

§ 31. The unique Factorization Theorem for polynomials over a Field. We shall now prove that every polynomial over a field can be factored uniquely into irreducible factors. Before stating the main factorization theorem, we shall give two preliminary theorems that are needed for its proof.

Theorem 1. *Let $f(x)$, $g(x)$, and $h(x)$ be polynomials in $F[x]$ for a field F. If $f(x) \mid g(x)\, h(x)$ and the greatest common divisor of $f(x)$ and $g(x)$ is 1, then $f(x) \mid h(x)$.*

Proof. If the greatest common divisor of $f(x)$ and $g(x)$ is 1, then by theorem of § 29 there exist polynomials $m(x)$ and $n(x) \in F[x]$ such that $1 = m(x) f(x) + n(x) g(x)$. Multiplying both members of this equation by $h(x)$, we get

$$h(x) = m(x) f(x) h(x) + n(x) g(x) h(x). \qquad \ldots(1)$$

But $f(x) \mid g(x)\, h(x)$, so there exists a polynomial $q(x) \in F[x]$ such that $g(x) h(x) = q(x) f(x)$.

Substituting this value of $g(x) h(x)$ in (1), we get

$$h(x) = m(x) f(x) h(x) + n(x) q(x) f(x)$$
$$= f(x) [m(x) h(x) + n(x) q(x)],$$

which shows that $f(x)$ is a divisor of $h(x)$.

Hence the theorem.

Theorem 2. *If $f(x)$ is an irreducible polynomial in $F[x]$ for a field F and $f(x) \mid g(x) h(x)$ where $g(x), h(x) \in F[x]$ then $f(x)$ divides at least one of $g(x)$ or $h(x)$.* **(Allahabad 1967)**

Proof. Suppose that $f(x)$ does not divide $g(x)$. Since $f(x)$ is prime therefore $f(x)$ does not divide $g(x)$ implies that $f(x)$ and $g(x)$ are relatively prime. Therefore the greatest common divisor of $f(x)$ and $g(x)$ is 1. Hence by theorem 1, we get that $f(x) \mid h(x)$.

Corollary. *If $f(x)$ is an irreducible polynomial in $F[x]$ for a field F, and if $f(x)$ divides the product $g_1(x) g_2(x) \ldots g_n(x)$ of polynomials in $F[x]$, then $f(x)$ divides $g_i(x)$ for some i, $1 \leqslant i \leqslant n$.*

This result follows immediately by repeated application of theorem 2.

The Unique Factorization Theorem for polynomials over a field. *Let $f(x)$ be a non-zero polynomial in $F[x]$, where F is a field. Then either $f(x)$ is a unit in $F[x]$ or $f(x)=ap_1(x)\,p_2(x)\ldots p_m(x)$, where each $p_i(x)$, $1 \leqslant i \leqslant m$, is an irreducible monic polynomial in $F[x]$ and $a \in F$ is the leading coefficient of $f(x)$. Further the factors $p_1(x), p_2(x), \ldots, p_m(x)$ are unique except for the order in which they appear.* **(Banaras 1969)**

Proof. We shall prove the theorem in two parts. First we shall prove that $f(x)$ can be factored as required, and then we shall show that the factors are unique.

Let $f(x)$ be a non-zero element of $F[x]$. Then either $f(x)$ is a unit in $F[x]$ *i.e.*, $\deg f(x)$ is 0 or $\deg f(x) > 0$. If $\deg f(x) > 0$, and the leading coefficient of $f(x)$ is a we are to prove that $f(x)$ can be expressed as a product of a and a finite number of irreducible monic polynomials in $F[x]$. The proof will be by induction on the degree of $f(x)$.

Suppose that $f(x)$ is of degree one. Let $f(x)=b+ax$ for $a, b \in F$ and $a \neq 0$. We can write $f(x)=a(a^{-1}b+x)$. Therefore the theorem holds in the case where $f(x)$ has degree one since $a^{-1}b+x$ is irreducible and monic.

Now assume, as the induction hypothesis, that every polynomial of degree less than n can be factored as stated in the theorem. Consider an arbitrary polynomial $f(x)$ of degree n having a as its leading coefficient. We can write $f(x)=af_1(x)$, where $f_1(x)=a^{-1}f(x)$ and $f_1(x)$ is monic. If $f(x)$ is irreducible, then $f_1(x)$ is also irreducible and the theorem holds If $f(x)$ is reducible, then it can be factored as $f(x)=g(x)\,h(x)$ where neither $g(x)$ nor $h(x)$ is a unit in $F[x]$ Now the degree of $f(x)$ is equal to the sum of the degrees of $g(x)$ and $h(x)$. Also $g(x)$ and $h(x)$ are not units in $F[x]$, so each of them must be of degree one or larger. Hence both $g(x)$ and $h(x)$ have degrees less than n. Therefore by our induction hypothesis we can write

$$g(x)=c\alpha_1(x)\,\alpha_2(x)\ldots\alpha_s(x),\ h(x)=d\beta_1(x)\,\beta_2(x)\ldots\beta_t(x)$$

where each $\alpha_i(x)$ and each $\beta_j(x)$ is monic and irreducible and where c and d are leading coefficients of $g(x)$ and $h(x)$ respectively. Thus

$$f(x)=cd\,\alpha_1(x)\,\alpha_2(x)\,\alpha_3(x)\ldots\alpha_s(x)\,\beta_1(x)\,\beta_2(x)\ldots\beta_t(x).$$

Since the leading coefficient of $f(x)$ is a, therefore we must have $a=cd$ because each $\alpha(x)$ and each $\beta(x)$ is monic. Therefore

$$f(x)=a\alpha_1(x)\,\alpha_2(x)\ldots\alpha_s(x)\,\beta_1(x)\,\beta_2(x)\ldots\beta_t(x).$$

The factorization of $f(x)$ satisfies the requirements of the theorem. Hence the theorem holds for all polynomials of degree n, and by the principle of induction, for all polynomials of arbitrary degree.

In order to prove that the factors are unique, let us suppose that $f(x)=ap_1(x)\,p_2(x)\ldots p_m(x)=aq_1(x)\,q_2(x)\ldots q_n(x)$ where each $p(x)$ and each $q(x)$ is irreducible and monic. Then we shall prove that $n=m$ and each $p(x)$ is equal to some $q(x)$ and each $q(x)$ is equal to some $p(x)$. From these two decompositions of $f(x)$, we have $\quad p_1(x)\,p_2(x)\ldots p_m(x)=q_1(x)\,q_2(x)\ldots q_n(x).$

Now $p_1(x)\mid p_1(x)\,p_2(x)\ldots p_m(x)$. Therefore

$$p_1(x)\mid q_1(x)\,q_2(x)\ldots q_n(x).$$

By Cor. to theorem 2 of this article $p_1(x)$ must divide at least one of $q_1(x), q_2(x),\ldots, q_n(x)$. Since $F[x]$ is a commutative ring, therefore without loss of generality we may suppose that $p_1(x)$ divides $q_1(x)$. But $p_1(x)$ and $q_1(x)$ are both irreducible polynomials in $F[x]$ and $p_1(x)\mid q_1(x)$. Therefore $p_1(x)$ and $q_1(x)$ must be associates and we have $q_1(x)=up_1(x)$ where u is a unit in $F[x]$ *i.e.*, u is a non-zero element of F. Since $q_1(x)$ and $p_1(x)$ are monic therefore u must be equal to 1 and we have $p_1(x)=q_1(x)$. Thus we have

$$p_1(x)\,p_2(x)\ldots p_m(x)=p_1(x)\,q_2(x)\ldots q_n(x).$$

Cancelling $0\neq p_1(x)$ from both sides, we get

$$p_2(x)\,p_3(x)\ldots p_m(x)=q_2(x)\,q_3(x)\ldots q_n(x) \qquad \ldots (1)$$

Now we can repeat the above argument on the relation (1) with $p_2(x)$. If $n > m$, then after m steps the left hand side becomes 1 while the right hand side reduces to a product of a certain number of $q(x)$ (the excess of n over m). But the $q(x)$ are irreducible polynomials so they are not units of $F[x]$ *i.e.*, they are not polynomials of zero degree.

So their product will be a polynomial of degree $\geqslant 1$. So it cannot be equal to 1. Therefore n cannot be greater than m. Then $n\leqslant m$. Similarly interchanging the roles of $p(x)$ and $q(x)$, we get $m\leqslant n$. Hence $m=n$.

Also in the above process we have shown that every $p(x)$ is equal to some $q(x)$ and conversely every $q(x)$ is equal to some $p(x)$. Hence the theorem has been completely established.

Thus we can say that *the ring of polynomials over a field is a unique factorization domain*

§ 32. Value of a polynomial at $x=c$.

Definition. *Let* $f(x)=a_0+a_1x+a_2x^2+\dots+a_nx^n$ *be a polynomial in* $F[x]$ *for an arbitrary field F and let c be an element of F. Then* $f(c)=a_0+a_1c+a_2c^2+\dots+a_nc^n$, *where the indicated addition and multiplication are the operations in F, is called the value of* $f(x)$ *at* $x=c$. Obviously $f(c)$ is an element of F.

Zeros of a polynomial Definition *If* $f(x)$ *is a polynomial in* $F[x]$ *for an arbitrary field F, and* $f(c)=0$ *for an element* $c \in F$, *then c is called a zero of* $f(x)$.

Polynomial equations and their roots Definition. *Let* $f(x)$ *be a polynomial of degree n over a field F. We say that* $f(x)=0$ *is an equation over the field F and n is the degree of the equation*

If c is a zero of the polynomial $f(x)$, *then c is a root of the equation* $f(x)=0$. A root of an equation is also called a **solution** of the equation.

Remainder Theorem. *If* $f(x) \in F[x]$ *and* $a \in F$, *for any field F, then* $f(a)$ *is the remainder when* $f(x)$ *is divided by* $(x-a)$.

Proof By division algorithm there exist polynomials $q(x)$ and $r(x)$ such that $f(x)=q(x)(x-a)+r(x)$, where either $r(x)=0$ or deg $r(x)$ is less than the degree of $x-a$. But the degree of $(x-a)$ is 1. Therefore $r(x)$ has degree 0 or no degree. Hence $r(x)$ is a constant polynomial *i.e.*, $r(x)$ is simply an element, say, r in F. Thus $f(x)=q(x)(x-a)+r$. Putting $x=a$ in this relation, we get $f(a)=q(a)(a-a)+r \Rightarrow f(a)=r$.

Cor. Factor Theorem. *If* $f(x) \in F[x]$ *and* $a \in F$, *for a field F, then* $x-a$ *divides* $f(x)$ *if and only if* $f(a)=0$.

(Meerut 1976; 84 P)

Proof. By remainder theorem, $f(a)$ is the remainder when $f(x)$ is divided by $(x-a)$. Therefore if $f(a)=0$, then $(x-a)$ divides $f(x)$.

Conversely, if $f(x)$ is divisible by $(x-a)$ we get

$$f(x)=(x-a)\,q(x).$$

Putting $x=a$, we get $f(a)=(a-a)\,q(a)=0\,q(a)=0$.

Example. *Show that the polynomial* x^2+x+4 *is irreducible over F, the field of integers modulo* 11. **(Meerut 1982, 83 P, 84 P)**

Solution. The field F is $(\{0, 1, \dots, 10\}, +_{11}, \times_{11})$.

Let $f(x)=x^2+x+4$.

If $a \in F$, then by a^n we shall mean $a \times_{11} a \times_{11} a \times_{11} a \ldots$ upto n times.

Now $f(0)=0^2+_{11} 0+_{11} 4=4$, $f(1)=1^2+_{11} 1+_{11} 4=6$, $f(2)=2^2+_{11} 2+_{11} 4=10$, $f(3)=3^2+_{11} 3+_{11} 4=5$, $f(4)=2$, $f(5)=1$, $f(6)=6^2+_{11} 6+_{11} 4=2$, $f(7)=5$, $f(8)=10$, $f(9)=6$, $f(10)=4$.

Since $f(a) \neq 0 \ \forall \ a \in F$, therefore by factor theorem $x-a$ does not divide $f(x) \ \forall \ a \in F$. Therefore $f(x)$ has no proper divisors in $F[x]$. Hence $f(x)$ is irreducible over F.

§ 33. **Prime fields. Definition.** *A field is said to be prime if it has no subfield other than itself.*

(Andhra 1977; Nagarjuna 79, 80; Delhi 70; Meerut 80, 81)

The field of rational numbers is a prime field while the field of real numbers is not a prime field The field $\mathbf{I}_p$ is prime for each prime integer p. **(Andhra 1977)**

Theorem 1. *Show that every prime field of characteristic* 0 *is isomorphic to the field of rational numbers.*

(Nagarjuna 1980; Kurushetra 70; Meerut 76, 81; Andhra 77; Sagar 66)

Proof. Let F be a prime field of characteristic 0. For the sake of convenience let us denote the unity element (multiplicative identity) of F by e. Since F is of characteristic 0, therefore for any integer n, we have $ne=0$ (zero element of F) if and only if $n=0$.

Here ne is an integral multiple of the element e of F. [For the def. of integral multiple see § 3 page 256 of this chapter on rings]. We have $ne \in F$. Consider a subset F' of F defined as

$F'=\{me/ne : m, n \in$ the set of integers $\mathbf{I}$ with $n \neq 0\}$.

Since $n \neq 0 \Rightarrow ne \neq 0$, therefore ne is an inversible element of F. So $me/ne=(me)(ne)^{-1}$ is definitely an element of F. We claim that F' is a subfield of F.

Let $\dfrac{m_1e}{n_1e}, \dfrac{m_2e}{n_2e}$ be any two elements of F'. Here $m_1, n_1, m_2, n_2 \in \mathbf{I}$ with $n_1 \neq 0$ and $n_2 \neq 0$. We have

$$\frac{m_1e}{n_1e}-\frac{m_2e}{n_2e}=\frac{(m_1e)(n_2e)-(n_1e)(m_2e)}{(n_1e)(n_2e)}=\frac{(m_1n_2)e^2-(n_1m_2)e^2}{(n_1n_2)e^2}$$

$$=\frac{(m_1n_2)e-(n_1m_2)e}{(n_1n_2)e} \quad [\because \ e^2=e]$$

$$=\frac{(m_1n_2-n_1m_2)e}{(n_1n_2)e} \in F' \text{ since } 0 \neq n_1n_2 \in \mathbf{I}.$$

Again let $\frac{m_1e}{n_1e}$ be any element of F' and $\frac{m_2e}{n_2e}$ be any non-zero element of F'. Then $m_1, n_1 \in \mathbf{I}$ with $n_1 \neq 0$. Also $m_2, n_2 \in \mathbf{I}$ with $m_2 \neq 0$, $n_2 \neq 0$. We have

$$\frac{m_1e}{n_1e}\left(\frac{m_2e}{n_2e}\right)^{-1}=\frac{m_1e}{n_1e}\,\frac{n_2e}{m_2e}=\frac{(m_1e)\,(n_2e)}{(n_1e)\,(m_2e)}=\frac{(m_1n_2)\,e^2}{(n_1m_2)\,e^2}$$

$$=\frac{(m_1n_2)\,e}{(n_1m_2)\,e}\in F' \text{ since } \neq n_1m_2 \in \mathbf{I}.$$

Therefore F' is a subfield of F. But F can have no proper subfield because F is a prime field. Therefore we must have $F'=F$.

Thus $F=\{me/ne : m, n \in \mathbf{I} \text{ with } n \neq 0\}$. If Q is the field of rational numbers, then $\mathbf{Q}=\{m/n : m, n \in \mathbf{I} \text{ with } n \neq 0\}$.

Let f be a mapping from F into Q defined as

$$f\,(me/ne)=m/n \;\forall\; m, n \in \mathbf{I} \text{ with } n \neq 0.$$

f is well-defined. We have $\frac{m_1e}{n_1e}=\frac{m_2e}{n_2e}$

$\Rightarrow (m_1e)\,(n_2e)=(n_1e)\,(m_2e) \Rightarrow (m_1n_2)\,e^2=(n_1m_2)\,e^2$

$\Rightarrow (m_1n_2)\,e=(n_1m_2)\,e$

$\Rightarrow (m_1n_2-n_1m_2)\,e=0 \Rightarrow m_1n_2-n_1m_2=0$

$\Rightarrow \frac{m_1}{n_1}=\frac{m_2}{n_2} \Rightarrow f\left(\frac{m_1e}{n_1e}\right)=f\left(\frac{m_2e}{n_2e}\right).$

$\therefore$ the mapping f is well-defined.

f is one-one. We have $f\left(\frac{m_1e}{n_1e}\right)=f\left(\frac{m_2e}{n_2e}\right)$

$\Rightarrow \frac{m_1}{n_1}=\frac{m_2}{n_2} \Rightarrow m_1n_2=n_1m_2 \Rightarrow (m_1n_2)e=(n_1m_2)e \Rightarrow (m_1n_2)e^2=(n_1m_2)e^2$

$\Rightarrow (m_1e)\,(n_2e)=(n_1e)\,(m_2e) \Rightarrow \frac{m_1e}{n_1e}=\frac{m_2e}{n_2e} \Rightarrow f$ is one-one

f is onto. Let m/n be any element of Q. Then $me/ne \in F$ and is such that $f\,(me/ne)=m/n$. Therefore f is onto.

f preserves compositions. We have

$$f\left(\frac{m_1e}{n_1e}+\frac{m_2e}{n_2e}\right)=f\left[\frac{(m_1e)\,(n_2e)+(n_1e)\,(m_2e)}{(n_1e)\,(n_2e)}\right]=f\left[\frac{(m_1n_2+n_1m_2)\;e}{(n_1n_2)\,e}\right]$$

$$=\frac{m_1n_2+n_1m_2}{n_1n_2}=\frac{m_1}{n_1}+\frac{m_2}{n_2}=f\left(\frac{m_1e}{n_1e}\right)+f\left(\frac{m_2e}{n_2e}\right).$$

Also $$f\left(\frac{m_1e}{n_1e}\,\frac{m_2e}{n_2e}\right)=f\left[\frac{(m_1m_2)\,e^2}{(n_1n_2)\,e^2}\right]=f\left[\frac{(m_1m_2)\;e}{(n_1n_2)\;e}\right]$$

$$=\frac{m_1m_2}{n_1n_2}=\frac{m_1}{n_1}\,\frac{m_2}{n_2}=f\left(\frac{m_1e}{n_1e}\right)f\left(\frac{m_2e}{n_2e}\right).$$

Hence $F \cong \mathbf{Q}$.

Theorem 2. *Every field of characteristic* 0 *contains a subfield isomorphic to the field of rational numbers* **(Nagarjuna 1979)**

Proof. Let F be any field of characteristic 0 and let e be the unity element of F. Since F is of characteristic 0, therefore for any integer n, we have $ne=0$ if and only if $n=0$.

Consider the subset F' of F defined as

$$F'=\{me/ne : m\in \mathbf{I},\ 0\neq n\in \mathbf{I}\}.$$

Now prove that F' is a subfield of F and $F' \cong \mathbf{Q}$ where $\mathbf{Q}$ is the field of rational numbers. Give the same proof as in theorem 1.

Theorem 3. *Every prime field of finite characteristic p is isomorphic to the field $\mathbf{I}_p$ of the residue classes of the set of integers modulo p.* **(Kurushetra 1970; Andhra 77; Nagarjuna 80)**

Proof. Let F be a prime field of finite characteristic p. Then p must be a prime number. The unit element e of F will be of order p regarded as an element of the additive group of F. The identity element of the additive group of F is the zero element of F. Therefore if n is any integer, then

$$ne=0 \text{ if and only if } p \text{ is a divisor of } n.$$

Consider a subset F' of F defined as

$$F'=\{ne : n\in \mathbf{I} \text{ where } \mathbf{I} \text{ is the set of integers}\}.$$

F' is a cyclic subgroup of the additive group of F.

Since F' is generated by e whose order is p, therefore F' contains p distinct elements. We claim that F' is a subfield of F. For this we shall prove that F' is an integral domain and we know that every finite integral domain is a field.

Let me, ne be any two elements of F'. Then

$$me-ne=(m-n)\, e \in F' \text{ since } m-n \in \mathbf{I}.$$

Also $(me)\,(ne)=(mn)\, e^2=(mn)\, e \in F'$ since $mn \in \mathbf{I}$.

Thus F' is a subring of F. Since F is without zero divisors, therefore F' is also without zero divisors. Therefore F' is a commutative ring without zero divisors. Therefore F' is an integral domain and so F' is a subfield of F. But F can have no proper subfield because F is a prime field. Therefore we must have

$$F=F'=\{ne : n\in \mathbf{I}\}.$$

Now we shall prove that $F \cong \mathbf{I}_p$.

Let f be a mapping from F into $\mathbf{I}_p$ defined as

$$f(ne)=\text{the residue class } [n],\ \forall\ n\in \mathbf{I}.$$

f is well defined. We have $ne=me$

$\Rightarrow (n-m)\, e=0 \Rightarrow p$ is a divisor of $n \quad m \Rightarrow n\equiv m \pmod p$

$\Rightarrow [n]=[m] \Rightarrow f(ne)=f(me) \Rightarrow f$ is well-defined.

f is one-one. We have $f(ne)=f(me)$
$\Rightarrow [n]=[m] \Rightarrow n-m$ is divisible by $p \Rightarrow (n-m)\, e=0$
$\Rightarrow ne=me \Rightarrow f$ is one-one.

f is onto. Let $[n]$ be any element of $\mathbf{I}_p$. Then $ne \in F$ and is such that $f(ne)=[n]$. Therefore f is onto.

f preserves compositions. We have

$f(me+ne)=f((m+n)\, e)=[m+n]=[m]+[n]=f(me)+f(ne)$.

Also $f((me)\,(ne))=f((mn)\, e)=[mn]=[m]\,[n]=f(me)\, f(ne)$.

Hence $F \cong \mathbf{I}_p$.

Theorem 4. *Let R be an integral domain with unity of finite characteristic p. Then R contains a subset isomorphic to the field $\mathbf{I}_p$ of the residue classes of the set of integers modulo p.*

(Banaras 1974; Nagarjuna 79)

Proof. Proceed as in theorem 3. If e is the unity element of R, then prove that $R'=\{ne : n \in \mathbf{I}\}$ is isomorphic to $\mathbf{I}_p$.

Theorem 5. *Let R be an integral domain with unity of characteristic 0. Then R contains a subset isomorphic to the integral domain of integers.* **(Banaras 1965)**

Proof. If e is the unity element of R, then prove that

$$R'=\{ne : n \in \mathbf{I}\}$$

is isomorphic to the integral domain $\mathbf{I}$ of integers. Show that the mapping f from R' into $\mathbf{I}$ defined as $f(ne)=n \ \forall\ n \in \mathbf{I}$ is an isomorphism of R' onto $\mathbf{I}$.

§ 34. The ring of endomorphisms of an abelian group.

Endomorphism of a group. Definition. *Let G be a group, the composition being denoted multiplicatively. A mapping f of G into itself is called an endomorphism of G if*

$$f(ab)=f(a)\, f(b) \ \forall\ a, b \in G.$$

Now suppose G is an abelian group, the operation in G being denoted additively. We recall that a mapping f of G into itself is an endomorphism of G, if

$$f(a+b)=f(a)+f(b) \ \forall\ a, b \in G.$$

Let R denote the set of all endomorphisms of an abelian group. We shall presently show that we can impose a ring structure on the set R by suitably defining the operations of addition and multiplication on it. This ring is called the ring of endomorphisms of an abelian group.

Theorem 1. *Let G be an abelian group (with the operation denoted additively) and let R be the set of all endomorphisms of G.*

Suppose addition and multiplication in R are defined as follows :

If $f, g \in R$, then for each $a \in G$,

$$(f+g)(a)=f(a)+g(a),\ (fg)(a)=f[g(a)].$$

Then under these operations R is a ring. **(Nagpur 1970; Andhra 79)**

Proof G is an additive abelian group and $R=\{f : f$ is an endomorphism of $G\}$. If $f, g \in R$, we have defined $f+g$ as follows :

$$(f+g)(a)=f(a)+g(a) \text{ for all } a \in G.$$

Since $f(a), g(a) \in G \Rightarrow f(a)+g(a) \in G$, therefore $f+g$ is a mapping of G into itself. We shall show that $f+g$ is also an endomorphism of G. For all $a, b \in G$, we have

$(f+g)(a+b)=f(a+b)+g(a+b)$ [by def. of $f+g$]

$=[f(a)+f(b)]+[g(a)+g(b)]$ [$\because$ f and g are endomorphisms]

$=[f(a)+g(a)]+[f(b)+g(b)]$ [$\because$ G is an abelian group]

$=(f+g)(a)+(f+g)(b).$ [by def. of $f+g$]

Therefore $f+g$ is an endomorphism of G. Thus $f \in R, g \in R \Rightarrow f+g \in R$ Therefore R is closed with respect to addition.

Further if $f, g \in R$, we have defined the multiplication fg in R as follows :

$$(fg)(a)=f[g(a)]\ \forall\ a \in G.$$

It should be noted that fg is nothing but the composite of the functions f and g. Obviously fg is a mapping of G into itself. We shall show that fg is also an endomorphism of G. For all $a, b \in G$, we have

$(fg)(a+b)=f[g(a+b)]$ [by def. of fg]

$=f[g(a)+g(b)]$ [$\because$ g is an endomorphism]

$=f[g(a)]+f[g(b)]$ [$\because$ f is also an endomorphism]

$=(fg)(a)+(fg)(b).$

Therefore fg is an endomorphism of G. Thus $f, g \in R \Rightarrow fg \in R$. Therefore R is closed with respect to multiplication.

Now we shall show that all the other ring postulates are also satisfied by these two operations on R. We should remember that if f and g are two mappings of a set G into itself, then $f=g$ if and only if $f(a)=g(a)\ \forall\ a \in G$.

Associativity of addition. For all $f, g, h \in R$ and for all $a \in G$ we have

$[(f+g)+h](a)=(f+g)(a)+h(a)$ [by def. of addition in R]

$=[f(a)+g(a)]+h(a)=f(a)+[g(a)+h(a)]$ [by associativity in G]

$=f(a)+(g+h)(a)=[f+(g+h)](a).$

Therefore by definition of equality of two functions, we have

$$(f+g)+h=f+(g+h).$$

Commutativity of addition. For all $a \in G$, we have

$$(f+g)(a)=f(a)+g(a)=g(a)+f(a)=(g+f)(a).$$

Therefore $\quad f+g=g+f.$

Existence of additive identity. Let 0 denote the identity element of the group G. Let us define a mapping $\hat{0}$ of G into itself by the rule $\hat{0}(a)=0, \forall a \in G$. For all $a, b \in G$, we have

$$\hat{0}(a+b)=0=0+0=\hat{0}(a)+\hat{0}(b).$$

Therefor $\hat{0}$ is an endomorphism of G *i.e.*, $\hat{0} \in R$. Now for all $f \in R$ and for all $a \in G$, we have

$$(\hat{0}+f)(a)=\hat{0}(a)+f(a)=0+f(a)=f(a).$$

Therefore $\hat{0}+f=f$. Thus $\hat{0} \in R$ is the additive identity.

Existence of the additive inverse. Let $f \in R$. Let us define a mapping $-f$ of G into itself by the rule

$$(-f)(a)=-f(a), \forall a \in G.$$

For all $a, b \in G$, we have

$$\begin{aligned}(-f)(a+b)&=-f(a+b)=-[f(a)+f(b)]\\&=[-f(a)]+[-f(b)]=(-f)(a)+(-f)(b).\end{aligned}$$

Therefore $-f$ is an endomorphism of G. Now for all $a \in G$, we have

$$(-f+f)(a)=(-f)(a)+f(a)=-f(a)+f(a)=0=\hat{0}(a).$$

Therefore $-f+f=\hat{0}$. Thus $-f \in R$ is the additive inverse of $f \in R$.

Associativity of multiplication. We know that composite or product of functions is an associative operation.

Distributive laws. For all $f, g, h \in R$ and for all $a \in G$, we have

$$\begin{aligned}[f(g+h)](a)&=f[(g+h)(a)]=f[g(a)+h(a)]=f[g(a)]+f[h(a)]\\&=(fg)(a)+(fh)(a)=[fg+fh](a).\end{aligned}$$

Therefore $f(g+h)=fg+fh$. Similarly we can show that

$$(g+h)f=gf+hf.$$

Thus R, has a ring structure for the addition and multiplication compositions as defined.

Note. The ring of endomorphisms of an abelian group G is a ring with unity. If 1 denotes the identity mapping of G *i.e.*, if $1 : G \to G$ such that $1\,(a)=a,\ \forall\ a \in G$, then 1 is the unit element of this ring. Obviously 1 is an endomorphism of G and we have $1\,f=f=f\,1\ \forall\ f \in R$. The ring R may not be commutative and may have zero divisors.

Theorem 2. *Every ring with unity is isomorphic to a ring of endomorphisms of an abelian group.* (Nagarjuna 1979)

Proof. Let R be a ring with unity element 1. The additive group of R is an abelian group. Let S denote the ring of endomorphisms of the abelian group R.

If $a \in R$, let f_a denote the mapping of R into itself defined by the rule $f_a\,(x)=ax\ \forall\ x \in R$.

Obviously f_a is an endomorphism of the additive group of R. For, if $x, y \in R$, we have

$$f_a\,(x+y)=a\,(x+y)=ax+ay=f_a(x)+f_a(y).$$

Thus f_a is an endomorphism of the additive group of R. Let $T=\{f_a : a \in R\}$. Then $T \subseteq S$. We shall show that T is a subring of S. First we shall show that

$$f_{a+b}=f_a+f_b,\ f_{ab}=f_a\,f_b,\ f_{-a}=-f_a.$$

Now for all $x \in R$, we have

$f_{a+b}(x)=(a+b)\,x=ax+bx=f_a(x)+f_b(x)=(f_a+f_b)\,(x).$

Therefore $f_{a+b}=f_a+f_b$.

Also $f_{ab}(x)=(ab)\,x=a(bx)=a\,[f_b(x)]=f_a\,[f_b(x)]=(f_a\,f_b)\,(x).$

Therefore $f_{ab}=f_a\,f_b$.

Further $f_{-a}\,(x)=(-a)\,x=-(ax)=-[f_a(x)]=(-f_a)\,(x).$

Therefore $f_{-a}=-f_a$.

Now let f_a, f_b be any two elements of T. We have $f_a-f_b=f_a+(-f_b)=f_a+f_{-b}=f_{a+(-b)}=f_{a-b}$. Since $a-b \in R$, therefore $f_{a-b} \in T$. Also $f_a f_b=f_{ab} \in T$ since $ab \in R$. Thus $f_a, f_b \in T \Rightarrow f_a-f_b \in T$ and $f_a f_b \in T$. Therefore T is a subring of S.

Now we shall show that the ring R is isomorphic to the ring T. Let $\phi : R \to T$ such that $\phi(a)=f_a\ \forall\ a \in R$.

ϕ is one one. If $a, b \in R$, then

$\phi(a)=\phi(b) \Rightarrow f_a=f_b \Rightarrow f_a(x)=f_b(x)\ \forall\ x \in R$
$\Rightarrow ax=bx\ \forall\ x \in R$
$\Rightarrow a1=b1$ $[\because\ 1 \in R]$
$\Rightarrow a=b.$

Therefore ϕ is 1-1.

ϕ is onto. Let $f_a \in T$. Then $a \in R$ and we have $\phi(a)=f_a$. Therefore ϕ is onto.

ϕ **preserves compositions in** R **and** T. Let $a, b \in R$. Then

$$\phi(a+b)=f_{a+b}=f_a+f_b=\phi(a)+\phi(b)$$

and

$$\phi(ab)=f_{ab}=f_a f_b=\phi(a)\,\phi(b).$$

Hence ϕ is an isomorphism of the ring R onto the ring T and therefore $R \cong T$.

Exercises

1. Show that if a ring R has no zero divisors, then the ring $R[x]$ has also no zero divisors. **(Meerut 1976)**

2. If p is a prime integer, show that it need not be a prime Gaussian integer.

3. Show that the polynomial x^2-9 is reducible over the ring of integers modulo 11.

4. Resolve x^4+4 into factors over the field

$(\{0, 1, 2, 3, 4\}, +_5, \times_5)$. **(Meerut 1980)**

Ans. $x^4+4=(x+1)(x+2)(x+3)(x+4)$.

5. Resolve x^2+1 into factors over the field Z_5.

(Meerut 1981, 82P, 83)

Ans. $x^2+1=(x+2)(x+3)$.

6. Find the solution of the equation $3x=2$ in the field $(Z_7, +_7, \times_7)$. **(Meerut 1981)**

Ans. $x=3$ because $3 \times_7 3=2$.

7. Show that $f(x)=x^3+8x-2$ is irreducible over **Q**. Is it irreducible over reals ? Give reasons for your answer. **(Meerut 1980)**

8. Let $f(x)=2x^4+3x^3+2$ and $g(x)=3x^5+4x^3+2x+3$ be two polynomials over the field

$$Z_5=(\{0, 1, 2, 3, 4\}, +_5, \times_5).$$

Determine (i) $(d/dx) f(x)$, (ii) $f(x).g(x)$. **(Meerut 1983)**

Sol. (i) We have

$$\frac{d}{dx} f(x)=4\,(2)\,x^3+3\,(3)\,x^2$$

$$=(2+_5 2+_5 2+_5 2)\,x^3+(3+_5 3+_5 3)\,x^2$$

$$=3x^3+4x^2.$$

(ii) We have $f(x)\, g(x)$

$$=(2+3x^3+2x^4)\,(3+2x^2+4x^3+3x^5)$$

$$=(2\times_5 3)+(2\times_5 2)\,x^2+[(2\times_5 4)+_5 (3\times_5 3)\ \ x^3+(2\times_5 3)\,x^4$$
$$+[(2\times_5 3)+_5 (3\times_5 2)]\,x^5+[(3\times_5 4)+_5 (2\times_5 2)]\,x^6$$
$$+(2\times_5 4)\,x^7+(3\times_5 3)\,x^8+(2\times_5 3)\,x^9$$

$$=1+4x^2+2x^3+x^4+2x^5+x^6+3x^7+4x^8+x^9.$$

9. If $f(x)=3x^7+2x+3$, $g(x)=5x^3+2x+6$ be two polynomials over the field $Z_7=(\{0, 1, 2, 3, 4, 5, 6\}, +_7, \times_7)$, determine
(i) $(d/dx)\, f(x)$, (ii) $f(x).g(x)$, and (iii) $f(x)+g(x)$.

(Meerut 1981)

Ans. (i) 2. (ii) $4+4x+4x^2+x^3+3x^4+4x^7+6x^8+x^{10}$.
(iii) $3x^7+5x^3+4x+2$.

10. Let $f(x)=x^6+3x^5+4x^2-3x+2$ and $g(x)=x^2+2x-3$ be in $Z_7[x]$. Find

(i) Sum and product of $f(x)$ and $g(x)$ in $Z_7[x]$.

(ii) Two polynomials $q(x)$ and $r(x)$ in $Z_7[x]$ such that $f(x)=q(x)\, g(x)+r(x)$ with degree of $r(x) < 2$. **(Meerut 1980 S)**

Ans. (i) $f(x)+g(x)=x^6+3x^5+5x^2+6x+6$;
$f(x)\ g(x)=1+6x+.\,x^2+5x^3+4x^4+5x^5+3x^6+5x^7+x^8$.
Note that in Z_7, we have $-3=4$, $-1=6$ etc.
(ii) $q(x)=x^4+x^3+x^2+x+5$, $r(x)=4x+3$.

11. Define a prime field. Prove that the field of rational numbers is a prime field. Give an example of a field which is not a prime field. **(Nagarjuna 1980)**

12. Prove that the field I_p is prime for each prime integer p.

13. Prove that the field of rational numbers and the field $J/(p)$ of residue classes modulo a prime p are the only prime fields apart from isomorphism. **(Nagarjuna 1980)**

5

Rings (Continued)

§ 1. Quotient Rings or Rings of Residue Classes. Suppose R is an arbitrary ring and S is an ideal (two sided ideal) in R. Then S is a subgroup of the additive abelian group of R. We can form the cosets (right as well as left) of S in R. Since R is an abelian additive group, therefore if $a \in R$, then the right coset $S+a$ will be equal to the corresponding left coset $a+S$. Thus we shall call $S+a$ as simply a coset of S in R. We remember from our study of cosets in group theory, that if $a, b \in R$, then

$$S+a=S+b \Leftrightarrow a-b \in S.$$

The cosets of S in R are called the residue classes of S in R. We denote the set of all residue classes of S in R by the symbol R/S. Thus $R/S=\{S+a : a \in R\}$.

We shall now impose a ring structure on the set R/S by defining addition and multiplication of residue classes.

Theorem. *If S is an ideal of a ring R, then the set*

$$R/S=\{S+a : a \in R\}$$

of all residue classes of S in R forms a ring for the two compositions in R/S defined as follows :

$(S+a)+(S+b)=S+(a+b)$ *[Addition of residue classes]*

$(S+a)(S+b)=S+ab$ *[Multiplication of residue classes]*

(I.A.S. 1973; Meerut 68; Raj. 78; Vikram 76;
Nagarjuna 78; Andhra 71; Kanpur 80)

Proof. Since $S+(a+b)$ and $S+ab$ are also residue classes of S in R, therefore R/S is closed with respect to addition and multiplication of residue classes. First of all, we shall show that both addition and multiplication in R/S are well defined. For this we are to show that if $S+a=S+a'$ and $S+b=S+b'$, then

$$(S+a)+(S+b)=(S+a')+(S+b')$$

and $$(S+a)(S+b)=(S+a')(S+b').$$

We have $$S+a=S+a' \Rightarrow a' \in S+a$$

and $$S+b=S+b' \Rightarrow b' \in S+b.$$

Therefore there exist $\alpha, \beta \in S$ such that $a' = \alpha + a$, $b' = \beta + b$.

Now $a' + b' = (\alpha + a) + (\beta + b) = (a + b) + (\alpha + \beta)$.

$\therefore \quad (a' + b') - (a + b) = \alpha + \beta \in S$.

$\therefore \quad S + (a' + b') = S + (a + b)$

$\Rightarrow (S + a') + (S + b') = (S + a) + (S + b)$.

Thus addition in R/S is well defined.

Again $a'b' = (\alpha + a)(\beta + b) = \alpha\beta + \alpha b + a\beta + ab$

$= ab + \alpha\beta + \alpha b + a\beta$.

$\therefore \quad a'b' - ab = \alpha\beta + \alpha b + a\beta \in S$. [Since S is an ideal therefore $\alpha, \beta \in S$ and $a, b \in R \Rightarrow \alpha b \in S$, $a\beta \in S$, $\alpha\beta \in S$ and finally

$\alpha\beta + \alpha b + a\beta \in S$].

Now since $a'b' - ab \in S$, therefore $S + a'b' = S + ab$

$\Rightarrow (S + a')(S + b') = (S + a)(S + b)$.

Hence multiplication in R/S is also well defined.

Associativity of addition in R/S. We have

$(S + a) + [(S + b) + (S + c)] = (S + a) + [S + (b + c)]$

$= S + [a + (b + c)] = S + [(a + b) + c] = [S + (a + b)] + (S + c)$

$= [(S + a) + (S + b)] + (S + c)$.

Commutativity of addition in R/S. We have

$(S + a) + (S + b) = S + (a + b) = S + (b + a) = (S + b) + (S + a)$.

Existence of additive identity. We have $S = S + 0 \in R/S$. If $S + a \in R/S$, then $(S + 0) + (S + a) = S + (0 + a) = S + a$.

$\therefore \quad S$ is the additive identity.

Existence of additive inverse. Let $S + a \in R/S$.

Then $S + (-a) \in R/S$. Also we have

$[S + (-a)] + [S + a] = S + [(-a) + a] = S + 0 = S$.

$\therefore \quad S + (-a)$ or $S - a$ is the additive inverse of $S + a$.

Associativity of multiplication. We have

$[(S + a)(S + b)](S + c) = (S + ab)(S + c) = S + (ab)c$

$= S + a(bc) = (S + a)(S + bc) = (S + a)[(S + b)(S + c)]$.

Distributivity of multiplication with respect to addition. We have

$(S + a)[(S + b) + (S + c)] = (S + a)[S + (b + c)]$

$= S + a(b + c) = S + (ab + ac) = (S + ab) + (S + ac)$

$= (S + a)(S + b) + (S + a)(S + c)$.

Similarly, we can prove that

$[(S + b) + (S + c)](S + a) = (S + b)(S + a) + (S + c)(S + a)$.

Hence R/S is a ring with respect to the two compositions. The residue class $S + 0$ or S is the zero element of this ring.

Note. The students should not confuse that by the multiplication $(S+a)(S+b)$ of residue classes we mean the totality of elements obtained on multiplying the elements of $S+a$ with those of $S+b$. This multiplication of residue classes is a new composition which we have defined in the set R/S.

Howerver the addition $(S+a)+(S+b)$ of residue classes as defined by us coincides with the totality of elements obtained on adding elements of $S+a$ to the elements of $S+b$ as can be easily seen :

$$(S+a)+(S+b)=S+(a+S)+b=S+(S+a)+b$$
$$=(S+S)+(a+b)=S+(a+b).$$

Ex. *If R/S is a ring of residue classes of S in R, prove that*

(*i*) *If R is commutative, so also is R/S.*

(*ii*) *If R has a unity element 1 so also has R/S, namely $S+1$.*

(Poona 1973)

Solution. (i) Suppose R is a commutative ring. Let $S+a$, $S+b$ be any two elements of R/S. Then $a, b \in R$ and $ab=ba$.

We have $(S+a)(S+b)=S+ab=S+ba=(S+b)(S+a)$.

$\therefore$ R/S is also a commutative ring.

(ii) Suppose R is a ring with unit element 1. Then $S+1 \in R/S$. If $S+a$ is any element of R/S, we have

$$(S+1)(S+a)=S+(1a)=S+a$$

and $$(S+a)(S+1)=S+(a1)=S+a.$$

$\therefore$ $S+1$ is the unit element of R/S.

§ 2. Homomorphism of rings.

Definition. Homomorphism into. *A mapping f from a ring R into a ring R' is said to be a homomorphism of R into R' if*

(*i*) $f(a+b)=f(a)+f(b)$ $\forall$ $a, b \in R$

(*ii*) $f(ab)=f(a)f(b)$ *for all* $a, b \in R$.

(Delhi 1970; Allahabad 80; Kumayun 77)

Homomorphism onto. *A mapping f from a ring R onto a ring R is said to be a homomorphism of R onto R' if*

(*i*) $f(a+b)=f(a)+f(b)$ $\forall$ $a, b \in R$.

(*ii*) $f(ab)=f(a)f(b)$ *for all* $a, b \in R$.

Also then R' is said to be a homomorphic image of R.

Theorem 1. *If f is a homomorphism of a ring R into a ring R', then*

(*i*) $f(0)=0'$, *where* 0 *is the zero element of the ring R and $0'$ is the zero element of R'.* **(Kumayun 1977; Allahabad 80)**

(*ii*) $f(-a)=-f(a)$ $\forall$ $a \in R$. **(Utkal 1969; Allahabad 80; Kumayun 77)**

Proof. (i) Let $a \in R$. Then $f(a) \in R'$. We have

$$f(a)+0'=f(a) \quad [\because \ 0' \text{ is the additive identity of } R']$$
$$=f(a+0)=f(a)+f(0).$$

Now R' is a group with respect to addition. Therefore

$$f(a)+0'=f(a)+f(0)$$
$$\Rightarrow 0'=f(0). \qquad \text{[by left cancellation law].}$$

(ii) Let a be any element of R. Then $-a \in R$.

We have $\quad 0'=f(0)=f[a+(-a)]=f(a)+f(-a)$.

$\therefore$ $f(-a)$ is the additive inverse of $f(a)$ in the ring R'. Thus $f(-a)=-f(a)$.

Theorem 2. *Let ϕ be a homomorphic mapping of a ring R into a ring R'. Let S' be the homomorphic image of R in R'. Then S' is a subring of R'.* **(Nagarjuna 1978; Karnatak 77)**

Proof. Since S' is the image of R in R' under the mapping ϕ, therefore $\quad \phi(R)=S' \subseteq R'$.

Let a', b' be any two elements of S'. Since $S'=\phi(R)$, therefore there exist elements $a, b \in R$ such that $\phi(a)=a'$, $\phi(b)=b'$.

We have $\quad a'-b'=\phi(a)-\phi(b)=\phi(a-b)$.

$[\because \ \phi$ is a homomorphism$]$

Now $a-b \in R$ is such that $a'-b'=\phi\,(a-b)$. Therefore

$$a'-b' \in S'.$$

Further $\quad a'b'=\phi(a)\,\phi(b)=\phi(ab) \in S'$, since $ab \in R$.

Thus $\quad a', b' \in S' \Rightarrow a'-b \in S'$ and $a'b' \in S'$.

Therefore S' is a subring of R'.

§ 3. Kernel of a ring homomorphism.

Definition. *If f is a homomorphism of a ring R into a ring R', then the set S of all those elements of R which are mapped onto the zero element of R' is called the kernel of the homomorphism f.*

Thus if f is a homomorphism of R into R', then S is the kernel of f if $S=\{x \in R : f(x)=0'$ where $0'$ is the zero element of $R'\}$.

Theorem . *If f is a homomorphism of a ring R into a ring R' with kernel S, then S is an ideal of R.*

(Andhra 1977; Karnatak 77; Banaras 70; Nagarjuna 80; Allahabad 80; Meerut 71, 78, 79, 81)

Proof. Let f be a homomorphism of a ring R into a ring R'. Let $0, 0'$ be the zero elements of R, R' respectively. Let S be the kernel of f. Then $S=\{x \in R : f(x)=0'\}$.

Since $f(0)=0'$, therefore at least $0 \in S$. Thus S is not empty.

Let $a, b \in S$. Then $f(a)=0', f(b)=0'$.

We have $f(a-b)=f[a+(-b)]=f(a)+f(-b)$

$$=f(a)-f(b)=0'-0'=0'.$$

$$\therefore \quad a-b \in S.$$

Also if r be any element of R, then

$$f(ar)=f(a)\,f(r)=0'\,f(r)=0'$$

and $$f(ra)=f(r)\,f(a)=f(r)\,0'=0'.$$

$\therefore \quad ar \in S,\ ra \in S.$

Thus $a, b \in S,\ r \in R \Rightarrow (a-b) \in S,\ ar \in S,\ ra \in S.$

$\therefore \quad S$ is an ideal of R.

Theorem 2. *The homomorphism ϕ of a ring R into a ring R' is an isomorphism of R into R' if and only if $I(\phi)=(0)$, where $I(\phi)$ denotes the kernel of ϕ.* **(Vikram 1975)**

Proof. Let ϕ be a homomorphism of a ring R into a ring R'. Let 0, $0'$ be the zero elements of R, R' respectively. Let $S=I(\phi)$ be the kernel of ϕ. Then S is an ideal of R and

$$S=\{a \in R : \phi(a)=0'\}.$$

Suppose ϕ is an isomorphism of R into R'. Then ϕ is one-one. Let $a \in S$. Then

$\phi(a)=0'$ [by def. of kernel]

$\Rightarrow \phi(a)=\phi(0)$ [$\because \ \phi(0)=0'$]

$\Rightarrow a=0.$ [$\because \ \phi$ is one-one]

Thus $a \in S \Rightarrow a=0$. In other words 0 is the only element of R which belongs to S. Therefore $S=(0)$.

Conversely suppose that $S=(0)$. Then to prove that ϕ is an isomorphism of R into R' *i.e.*, to prove that ϕ is one-one.

If $a, b \in R$, then $\phi(a)=\phi(b)$

$\Rightarrow \phi(a)-\phi(b)=0'$ [$\because \ \phi(a), \phi(b)$ are in the ring R']

$\Rightarrow \phi(a-b)=0'$ [$\because \ \phi$ is a homomorphism]

$\Rightarrow a-b \in S$ [by def. of kernel]

$\Rightarrow a-b=0$ [$\because \ S=(0)$]

$\Rightarrow a=b.$

$\therefore \quad \phi$ is one-one. Hence ϕ is an isomorphism of R into R'.

Theorem 3. *Suppose R is a ring, S an ideal of R. Let f be a mapping from R to R/S defined by $f(a)=S+a \ \forall \ a \in R$. Then f is a homomorphism of R onto R/S.* **(I A.S. 1973; Vikram 75; Kanpur 80; Nagarjuna 78; Karnatak 77; Andhra 77)**

Proof. Consider the mapping of $f : R \to R/S$ such that

$$f(a)=S+a \ \forall \ a \in R.$$

Let $S+x$ be any element of R/S. Then $x \in R$.

We have $f(x)=S+x$. Therefore the mapping f is onto R/S.

Let $a, b \in R$. Then

$$f(a+b)=S+(a+b)=(S+a)+(S+b)=f(a)+f(b).$$

Also $f(ab)=S+ab=(S+a)(S+b)=f(a)f(b)$.

$\therefore$ f is a homomorphism of R onto R/S.

Thus every quotient ring of a ring is a homomorphic image of the ring.

Theorem 4. Fundamental theorem on homomorphism of rings.

Every homomorphic image of a ring R is isomorphic to some residue class ring (quotient ring) thereof.

(Nagarjuna 1979, 80; Meerut 79, 83 P, 84)

Proof. Let R' be the homomorphic image of a ring R and f be the corresponding homomorphism. Then f is a homomorphism of R onto R'. Let S be the kernel of this homomorphism. Then S is an ideal of R. Therefore R/S is a ring of residue classes of R relative to S. We shall prove that $R/S \cong R'$.

If $a \in R$, then $S+a \in R/S$ and $f(a) \in R'$. Consider the mapping $\phi : R/S \to R'$ such that

$$\phi(S+a)=f(a) \ \forall \ a \in R.$$

First we shall show that the mapping ϕ is well defined *i.e.*, if $a, b \in R$ and $S+a=S+b$, then $\phi(S+a)=\phi(S+b)$.

We have $\quad S+a=S+b$

$\Rightarrow a-b \in S$

$\Rightarrow f(a-b)=0'$ $\qquad$ [*i.e.* zero element of R']

$\Rightarrow f[a+(-b)]=0' \Rightarrow f(a)+f(-b)=0' \Rightarrow f(a)-f(b)=0'$

$\Rightarrow f(a)=f(b) \Rightarrow \phi(S+a)=\phi(S+b)$.

$\therefore$ ϕ is well-defined.

ϕ is one-one. We have $\phi(S+a)=\phi(S+b)$

$\Rightarrow f(a)=f(b) \Rightarrow f(a)-f(b)=0'$

$\Rightarrow f(a)+f(-b)=0' \Rightarrow f(a-b)=0'$

$\Rightarrow a-b \in S$ $\qquad$ [$\because$ S is kernel of f]

$\Rightarrow S+a=S+b$.

$\therefore$ ϕ is one-one.

ϕ is onto R'. Let y be any element of R'. Then $y=f(a)$ for some $a \in R$ because f is onto R'. Now $S+a \in R/S$ and we have $\phi(S+a)=f(a)=y$. Therefore ϕ is onto R'.

Finally we have

$$\phi[(S+a)+(S+b)]=\phi[S+(a+b)]=f(a+b)$$
$$=f(a)+f(b)=\phi(S+a)+\phi(S+b).$$

Also $\phi[(S+a)(S+b)]=\phi(S+ab)=f(ab)=f(a)f(b)$

$$=[\phi(S+a)][\phi(S+b)].$$

$\therefore$ ϕ is an isomorphism of R/S onto R'.

Hence $$R/S \cong R'.$$

Solved Examples

Ex. 1. *Show that every homomorphic image of a commutative ring is commutative.*

Solution. Let R be a commutative ring. Let f be a homomorphic mapping of R onto a ring R'. Then R' is a homomorphic image of R.

Let a', b' be any two elements of R'. Then $f(a)=a'$, $f(b)=b'$ for some $a, b \in R$ because f is onto R'. We have

$$\begin{aligned} a'b' &= f(a)\, f(b) = f(ab) \\ &= f(ba) \qquad [\because R \text{ is commutative}] \\ &= f(b)\, f(a) = b'a'. \end{aligned}$$

$\therefore$ R' is a commutative ring.

Ex. 2. *If R is a ring with unit element 1 and ϕ is a homomorphism of R onto R' prove that $\phi(1)$ is the unit element of R'.*

(Allahabad 1980)

Solution. Since ϕ is a homomorphism of R onto R', therefore R' is a homomorphic image of R. If 1 is the unity element of R, then $\phi(1) \in R'$. Let a' be any element of R'. Then $a'=\phi(a)$ for some $a \in R$ since ϕ is onto R'. We have

$$\phi(1)\, a' = \phi(1)\, \phi(a) = \phi(1a) = \phi(a) = a'$$

and $$a'\, \phi(1) = \phi(a)\, \phi(1) = \phi(a1) = \phi(a) = a'.$$

$\therefore$ $\phi(1)$ is the unity element of R'.

Ex. 3. *If R is a ring with unit element 1 and ϕ is a homomorphism of R into an integral domain R' such that kernel of ϕ i.e., $I(\phi) \neq R$, then prove that $\phi(1)$ is the unit element of R'.*

Solution. ϕ is a homomorphism of a ring R into an integral domain R'. Then kernel of ϕ

$$= I(\phi) = \{x : x \in R \text{ and } \phi(x) = 0 \in R'\}.$$

Since $I(\phi) \neq R$, therefore there exists an element $a \in R$ such that $$\phi(a) \neq 0 \in R'.$$

We have $\phi(1)\, \phi(a) = \phi(1a) = \phi(a)$.

Now let b' be any element of R'. We have

$$\begin{aligned} &\phi(a)\, b' = \phi(a)\, b' \\ \Rightarrow\ &\phi(1)\, \phi(a)\, b' = \phi(a)\, b' \qquad [\because \phi(1)\, \phi(a) = \phi(a)] \\ \Rightarrow\ &\phi(a)\, [\phi(1)\, b'] = \phi(a)\, b' \end{aligned}$$

[$\because$ $\phi(1)$, $\phi(a) \in R'$ which, being an integral domain, is a commutative ring]

$$\Rightarrow \phi(a)\, [\phi(1)\, b'] - \phi(a)\, b' = 0$$

$\Rightarrow \phi(a)\,[\phi(1)\,b'-b']=0$

$\Rightarrow \phi(1)\,b'-b'=0$

$[\because \phi(a)\neq 0$ and R' is without zero divisors$]$

$\Rightarrow \phi(1)\,b'=b'=b'\,\phi(1).$ $[\because R'$ is a commutative ring$]$

Thus $\phi(1)\,b'=b'=b'\,\phi(1)\ \forall\ b'\in R'$.

$\therefore$ $\phi(1)$ is the unit element of R'.

Ex. 4. *Prove that any homomorphism of a field is either an isomorphism or takes each element into* 0. **(Nagarjuna 1979)**

Or

Show that a field has no proper homomorphic image.

Solution. Let ϕ be a homomorphism of a field F into a ring R. Let S be the kernel of ϕ. Then S is an ideal of the field F. We know that a field has no proper ideals. Therefore either $S=F$ or $S=(0)$.

If $S=F$, then by definition of kernel of ϕ, we have $\phi(x)=0$ $\forall\ x\in F$. Thus in this case ϕ takes each element of F into the zero element of R In other words in this case $\phi(F)$ is the zero subring of the ring R.

If $S=(0)$, then the kernel consists of zero element alone. So in this case ϕ is an isomorphism of F into R. [See theorem 2 on page 358]. Since the isomorphic image of a field is a field, therefore in this case $\phi(F)$ is a field isomorphic to the filed F.

§ 4. **Maximal Ideal.** *An ideal $S\neq R$ in a ring R is said to be a maximal ideal of R if whenever U is an ideal of R such that $S\subseteq U\subseteq R$, then either $R=U$ or $S=U$.*

(Banaras 1971; Meerut 80, 84 P; Karnatak 77; Raj. 78; Guru Nanak 82)

In other words an ideal S of a ring R is said to be *maximal ideal* if there exists no ideal properly contained in R which itself properly contains S *i.e.*, if it is impossible to find an ideal which lies between S and the full ring R. For example, in the ring of integers **I**, the ideal (6) is not maximal since it is properly contained in the ideal (3), which in turn is properly contained in **I**. On the other hand, (5) is a maximal ideal since the only ideal properly contatining (5) is **I** itself.

Ex. *If A be the ring of all functions from the set of reals to itself under pointwise operations and I, the subset of all functions vanishing at zero, prove that I is a maximal ideal of A.*

Theorem. *An ideal S of the ring of integers* **I** *is maximal if and only if S is generated by some prime integer.*

(Meerut 1979, 83 P, 84 P)

Proof. We know that every ideal of the ring of integers **I** is a principal ideal. Suppose S is an ideal of **I** generated by p so that $S=(p)$. Since p and $-p$ generate the same ideal, therefore we can take p as positive.

Now we are to prove that

(i) S is maximal if p is prime.

(ii) p is prime if S is maximal.

First we shall prove (i). Let p be a prime integer such that $(p)=S$. Let T be an ideal of **I** such that $S \subseteq T \subseteq \mathbf{I}$. Since T is also a principal ideal of **I**, let $T=(q)$ where q is some positive integer.

Now
$$\begin{aligned} S \subseteq T &\Rightarrow (p) \subseteq T \\ &\Rightarrow p \in T \\ &\Rightarrow p \in \{xq : x \in \mathbf{I}\} \\ &\Rightarrow p=rq \text{ for some positive integer } r. \end{aligned}$$

Since p is prime, therefore $q=1$ or $q=p$.

If $q=1$, we have $T=(q)=(1)=\mathbf{I}$

and if $q=p$, we have $T=(q)=(p)=S$.

Thus either $T=\mathbf{I}$ or $T=S$.

Hence (p) is a maximal ideal of **I**

Now we shall prove (ii). Let $(p)=S$ be a maximal ideal. We are to show that p is prime. Let us suppose that p is a composite integer.

Let $p=mn$, $m \neq 1$, $n \neq 1$.

It is obvious that $(p) \subseteq (m) \subseteq \mathbf{I}$.

But since (p) is a maximal ideal, therefore we have

$$\text{either } (m)=(p) \text{ or } (m)=\mathbf{I}.$$

If $(m)=\mathbf{I}$, then $m=1$, which is a contradiction.

If $(m)=(p)$, then m must be equal to lp for some integer l since each element of (p) is a multiple of p.

Therefore $p=mn=lpn=pln$. But $p \neq 0$, therefore $ln=1$. This gives $n=1$ which is again a contradiction.

Hence p must be a prime integer.

§ 5. Some more results on Ideals.

Theorem 1. *Let S_1, S_2 be ideals of a ring R and let*

$$S_1+S_2=\{s_1+s_2 : s_1 \in S_1,\ s_2 \in S_2\}.$$

Then S_1+S_2 is an ideal of R generated by $S_1 \cup S_2$. **(Vikram 1976)**

Proof. Let $a_1+a_2 \in S_1+S_2$, $b_1+b_2 \in S_1+S_2$. Then $a_1, b_1 \in S_1$ and $a_2, b_2 \in S_2$.

We have $(a_1+a_2)-(b_1+b_2)=(a_1-b_1)+(a_2-b_2)$.

Since S_1 is an ideal, therefore $a_1, b_1 \in S_1 \Rightarrow a_1-b_1 \in S_1$. Similarly $a_2-b_2 \in S_2$.

$$\therefore \quad (a_1-b_1)+(a_2-b_2) \in S_1+S_2.$$

$$\therefore \quad (a_1+a_2)-(b_1+b_2) \in S_1+S_2.$$

$\therefore$ S_1+S_2 is a subgroup of the additive group of R.

Let r be any element of R. Then

$r(a_1+a_2)=ra_1+ra_2$

$\in S_1+S_2$ since $r \in R$, $a_1 \in S_1 \Rightarrow ra_1 \in S_1$

and similarly $ra_2 \in S_2$.

Similarly $(a_1+a_2)\,r=a_1r+a_2\,r \in S_1+S_2$ since $a_1r \in S_1$, $a_2r \in S_2$.

Hence S_1+S_2 is an ideal of R.

Since $0 \in S_1$ and also $0 \in S_2$, therefore obviously

$$S_1 \subseteq S_1+S_2 \text{ and } S_2 \subseteq S_1+S_2.$$

$$\therefore \quad S_1 \cup S_2 \subseteq S_1+S_2.$$

Thus S_1+S_2 is an ideal of R containing $S_1 \cup S_2$.

Also if S is any ideal of R containing $S_1 \cup S_2$, then S must contain S_1+S_2. Thus S_1+S_2 is the smallest ideal of R containing $S_1 \cup S_2$ *i e.*, $S_1+S_2=(S_1 \cup S_2)$.

Theorem 2. *If an ideal U of a ring R contains a unit of R then $U=R$.*

Proof. Let R be a ring with unity element 1. Let u be a unit of R. Then u is an inversible element of R *i.e.*, u^{-1} exists. Let $u \in U$.

Since U is an ideal, therefore

$$u \in U,\ u^{-1} \in R \Rightarrow uu^{-1} \in U \Rightarrow 1 \in U.$$

Now let x be any element of R. Then

$$x \in R,\ 1 \in U \Rightarrow x1 \in U \Rightarrow x \in U.$$

$$\therefore \quad R \subseteq U.$$

Also $U \subseteq R$ as U is an ideal of R. Hence $R=U$.

Theorem 3. *Let R be a commutative ring with unity and a, b be two non-zero elements of R. Then*

$$(a)=(b) \text{ iff } a \mid b \text{ and } b \mid a.$$

Proof. Here $(a)=$the principal ideal of R generated by a

$=\{ax : x \in R\}$.

Similarly $(b)=$the principal ideal of R generated by b.

Let $(a)=(b)$.

Then $(a) \subseteq (b)$

$\Rightarrow a \in (b)$

$\Rightarrow a=rb$ for some $r \in R$

$\Rightarrow b \mid a$ *i.e.*, b is a divisor of a.

Similarly $(a)=(b)$

$\Rightarrow$ $(b) \subseteq (a) \Rightarrow b \in (a)$

$\Rightarrow$ $b=sa$ for some $s \in R \Rightarrow a \mid b$.

Thus $(a)=(b) \Rightarrow a \mid b$ and $b \mid a$.

Conversely let $a \mid b$ and $b \mid a$.

Now $a \mid b \Rightarrow b=pa$ for some $p \in R$. Let y be any element of (b). Then $y=ub$ for some $u \in R$

$=u\,(pa)=(up)\,a \in (a)$ since $up \in R$.

Thus $y \in (b) \Rightarrow y \in (a)$.

$\therefore$ $(b) \subseteq (a)$.

Thus $a \mid b \Rightarrow (b) \subseteq (a)$. Similarly $b \mid a \Rightarrow (a) \subseteq (b)$.

Consequently $a \mid b$, $b \mid a \Rightarrow (a)=(b)$.

Corollary. *Let R be an integral domain with unity and a, b be two non-zero elements of R. Then $(a)=(b)$ iff a and b are associates.*

Theorem 4. *An ideal S of a commutative ring R with unity is maximal if and only if the residue class ring R/S is a field.*

(I.A.S. 1971; Madras 83; Karnatak 77; Marathwada 72; G.N.D.U. Amritsar 82; Meerut 79, 83 P)

Proof. Since R is a commutative ring with unity, therefore R/S is also a commutative ring with unity. The zero element of the ring R/S is S and the unity element is the coset $S+1$ where 1 is the unity element of R.

Let the ideal S be maximal. Then to prove that R/S is a field.

Let $S+b$ be any non-zero element of R/S. Then $S+b \neq S$ *i.e.*, $b \notin S$. To prove that $S+b$ is inversible.

If (b) is the principal ideal of R generated by b, then $S+(b)$ is also an ideal of R. Since $b \notin S$, therefore the ideal S is properly contained in $S+(b)$. But S is a maximal ideal of R. Hence we must have $S+(b)=R$.

Since $1 \in R$, therefore we must obtain 1 on adding an element of S to an element of (b). Therefore there exists an element $a \in S$ and $\alpha \in R$ such that

$a+\alpha b=1$ [Note that $(b)=\{\alpha b : \alpha \in R\}$]

$\therefore$ $1-\alpha b=a \in S$.

Consequently $S+1=S+\alpha b=(S+\alpha)(S+b)$.

$\therefore$ $S+\alpha=(S+b)^{-1}$. Thus $S+b$ is inversible.

$\therefore$ R/S is field.

Conversely, let S be an ideal of R such that R/S is a field. We shall prove that S is a maximal ideal of R.

Let S' be an ideal of R properly containing S *i.e.*, $S\subseteq S'$ and $S\neq S'$. Then S will be maximal if $S'=R$. The elements of R contained in S already belong to S' since $S\subseteq S'$. Therefore R will be a subset of S' if every element α of R not contained in S also belongs to S'. If $\alpha\in R$ is such that $\alpha\notin S$, then $S+\alpha\neq S$ *i.e.*, $S+\alpha$ is a non-zero element of R/S. Also S' properly contains S. Therefore there exists an element β of S' not contained in S so that $S+\beta$ is also a non-zero element of R/S. Now the non-zero elements of R/S form a group with respect to multiplication because R/S is a field. Therefore there exists a non-zero element $S+y$ of R/S such that

$$(S+y)\,(S+\beta)=S+\alpha.$$

[We may take $S+y=(S+\alpha)\,(S+\beta)^{-1}$].

Now $(S+y)\,(S+\beta)=S+\alpha$

$\Rightarrow S+y\beta=S+\alpha \Rightarrow y\beta-\alpha\in S\Rightarrow y\beta-\alpha\in S'.$ [$\because$ $S\subseteq S'$]

Now S' is an ideal. Therefore $y\in R,\ \beta\in S'\Rightarrow y\beta\in S'$. Again $y\beta\in S',\ y\beta-\alpha\in S'\Rightarrow y\beta-(y\beta-\alpha)\in S'$ *i.e.*, $\alpha\in S'$.

Thus $R\subseteq S'$. Also $S'\subseteq R$ as S' is an ideal of R.

$\therefore\ S'=R.$

Hence the theorem.

§ 6. Prime Ideals.

Prime Ideal. Definition. *Let R be a ring and S an ideal in R. Then S is said to be a prime ideal of R if* $ab\in S$, $a, b\in R$ *implies that either a or b is in S.* **(Gujrat 1971; Karnatak 77; Raj. 78; I.A.S. 75; Guru Nanak 75)**

For example, in the ring of integers **I**, the principal ideal (7) is prime. Obviously if *ab* is in (7), then *a* or *b* must be a multiple of 7. On the other hand, (6) is not a prime ideal in **I** since, in particular, $12=3\times 4$ is in (6), yet neither 3 nor 4 is an element of (6).

Theorem 1. *Let R be a commutative ring and S an ideal of R. Then the ring of residue classes R/S is an integral domain if and only if S is a prime ideal.* **(Gujrat 1970; Delhi 70; Indore 70)**

Proof. Let R be a commutative ring and S an ideal of R. Then $R/S=\{S+a : a\in R\}$.

Let $S+a$, $S+b$ be any two elements of R/S. Then $a, b\in R$.

We have $(S+a)\,(S+b)=S+ab$

$=S+ba$ [$\because$ R is a commutative ring]

$=(S+b)\,(S+a)$.

$\therefore$ R/S is a commutative ring.

Now let S be a prime ideal of R. Then we are to prove that R/S is an integral domain. For this we are to show that R/S is without zero divisors. The zero element of the ring R/S is the residue class S itself. Let $S+a$, $S+b$ be any two elements of R/S.

Then $(S+a)(S+b)=S$ (the zero element of R/S)

$\Rightarrow S+ab=S \Rightarrow ab \in S$

$\Rightarrow$ either a or b is in S, since S is a prime ideal

$\Rightarrow$ either $S+a=S$ or $S+b=S$ [Note that $a \in S \Leftrightarrow S+a=S$]

$\Rightarrow$ either $S+a$ or $S+b$ is the zero element of R/S.

$\therefore$ S/R is without zero divisors.

Since R/S is a commutative ring without zero divisors, therefor R/S is an integral domain.

Conversely, let R/S be an integral domain. Then we are to prove that S is a prime ideal of R. Let a, b be any two elements in R such that $ab \in S$. We have

$$ab \in S \Rightarrow S+ab=S \Rightarrow (S+a)(S+b)=S.$$

Since R/S is an integral domain, therefore it is without zero divisors. Therefore

$(S+a)(S+b)=S$ (the zero element of R/S)

$\Rightarrow$ either $S+a$ or $S+b$ is zero $\Rightarrow$ either $S+a=S$ or $S+b=S$

$\Rightarrow$ either $a \in S$ or $b \in S \Rightarrow S$ is a prime ideal.

This completes the proof of the theorem.

Note. If R is a ring with unity, then R/S is also a ring with unity. The residue class $S+1$ is the unity element of R/S. Therefore if we define an integral domain as a commutative ring with unity and without zero divisors, even then the above theorem will be true. But in that case R must be a commutative ring with unity.

Theorem *Let R be a commutative ring with unity. Then every maximal ideal of R is a prime ideal.*

(I.A.S. 1970; Karnatak 77; Gujrat 1971; Calicut 75)

Proof. R is a commutative ring with unit element. Let S be a maximal ideal of R. Then R/S is a field.

Now every field is an integral domain. Therefore R/S is also an integral domain. Hence by theorem 1, S is a prime ideal of R. This completes the proof of the theorem.

But it should be noted that the converse of the above theorem is not true *i.e.*, every prime ideal is not necessarily a maximal ideal.

Solved Examples

Ex. 1 *Let* R *be the field of real numbers and* S *the set of all* *those polynomials $f(x) \in \mathbf{R}[x]$ such that $f(0)=0=f(1)$. Prove that S is an ideal of $\mathbf{R}[x]$. Is the residue class ring $\mathbf{R}[x]/S$ an integral domain ? Give reasons for your answer.* **(Gujrat 1970)**

Solution. Let $f(x)$, $g(x)$ be any elements of S. Then

$$f(0)=0=f(1) \text{ and } g(0)=0=g(1).$$

Let $h(x)=f(x)-g(x)$. Then

$h(0)=f(0)-g(0)=0-0=0$ and $h(1)=f(1)-g(1)=0-0=0$.

Thus $h(0)=0=h(1)$. Therefore $h(x) \in S$.

Thus $f(x), g(x) \in S \Rightarrow h(x)=f(x)-g(x) \in S$.

Further let $f(x)$ be any element of S and $r(x)$ be any element of $\mathbf{R}[x]$. Then $f(0)=0=f(1)$, by definition of S.

Let $t(x)=r(x)\,f(x)=f(x)\,r(x)$ [$\because$ $\mathbf{R}[x]$ is a commutative ring]

Then $t(0)=r(0)\,f(0)=r(0).0=0$

and $t(1)=r(1)\,f(1)=r(1).0=0$.

$\therefore$ $t(x) \in S$.

Thus $r(x) \in \mathbf{R}[x], f(x) \in S \Rightarrow r(x) f(x) \in S$.

Hence S is an ideal of $\mathbf{R}[x]$.

Now we claim that S is not a prime ideal of $\mathbf{R}[x]$. Let $f(x)=x(x-1)$. Then $f(0)=0(0-1)=0$, and $f(1)=1(1-1)=0$.

Thus $f(x)=x(x-1)$ is an element of S.

Now let $p(x)=x$, $q(x)=x-1$.

We have $p(1)=1 \neq 0$. Therefore $p(x) \notin S$. Also $q(0)=0-1=-1 \neq 0$. Therefore $q(x) \notin S$. Thus $x(x-1) \in S$ while neither $x \in S$ nor $x-1 \in S$. Hence S is not a prime ideal of $\mathbf{R}[x]$.

Since S is not a prime ideal of $\mathbf{R}[x]$, therefore the residue class ring $\mathbf{R}[x]/S$ is not an integral domain.

Ex. 2. *Let R be the ring of all real valued continuous functions defined on the closed interval $[0, 1]$. Let*

$$M=\{f(x) \in R : f(\tfrac{1}{3})=0\}.$$

Show that M is a maximal ideal of R. **(Guru Nanak 1982)**

Solution. First of all we observe that M is non-empty because the real valued function $e(x)$ on $[0, 1]$ defined by

$$e(x)=0 \ \forall \ x \in [0, 1]$$

belongs to M.

Now let $f(x)$, $g(x)$ be any two elements of M. Then $f(\frac{1}{3})$ 0, $g(\frac{1}{3})=0$, by definition of M.

Let $h(x)=f(x)-g(x)$. Then $h(\frac{1}{3})=f(\frac{1}{3})-g(\frac{1}{3})=0-0=0$.

Therefore $h(x) \in M$.

Thus $f(x), g(x) \in M \Rightarrow h(x) = f(x) - g(x) \in M$.

Further let $f(x)$ be any element of M and $r(x)$ be any element of R. Then $f(\frac{1}{3}) = 0$, by definition of M.

Let $t(x) = r(x) f(x) = f(x) r(x)$. [$\because$ R is a commutative ring].

Then $t(\frac{1}{3}) = r(\frac{1}{3}) f(\frac{1}{3}) = r(\frac{1}{3}) . 0 = 0$. Therefore $t(x) \in M$.

Thus $r(x) \in R, f(x) \in M \Rightarrow r(x) f(x) \in M$.

Hence M is an ideal of R.

Clearly $M \neq R$ because $i(x) \in R$ given by $i(x) = 1 \; \forall \; x \in [0, 1]$ does not belong to M.

The ring R is with unity and the element $i(x)$ is its unity element.

Let N be an ideal of R properly containing M *i e.*, $M \subseteq N$ and $M \neq N$. Then M will be a maximal ideal of R if $N = R$, which will be so if the unity $i(x)$ of R belongs to N. Since M is a proper subset of N, therefore there exists $\lambda(x) \in N$ such that $\lambda(x) \notin M$. This means $\lambda(\frac{1}{3}) \neq 0$. Put $\lambda(\frac{1}{3}) = c$ where $c \neq 0$.

Let us define $\beta(x) \in R$ by $\beta(x) = c \; \forall \; x \in [0, 1]$. Now consider $\mu(x) \in R$ given by $\mu(x) = \lambda(x) - \beta(x)$.

We have $\qquad \mu(\frac{1}{3}) = \lambda(\frac{1}{3}) \quad \beta(\frac{1}{3}) = c - c = 0$

Therefore $\mu(x) \in M$ and so $\mu(x)$ also belongs to N because N is a super-set of M. Now N is an ideal of R and $\lambda(x)$, $\mu(x)$ are in N. Therefore $\lambda(x) - \mu(x) = \beta(x)$ is also an element of N.

Now define $\gamma(x) \in R$ by $\gamma(x) = 1/c \; \forall \; x \in [0, 1]$. Since N is an ideal of R, therefore $\gamma(x) \in R$ and $\beta(x) \in N \Rightarrow \gamma(x) \, \beta(x) \in N$. We shall show that $\gamma(x) \, \beta(x) = i(x)$.

For every $x \in [0, 1]$, we have

$$\gamma(x) \, \beta(x) = (1/c) \; c = 1.$$

Therefore $\gamma(x) \, \beta(x) = i(x)$, by the definition of $i(x)$.

Thus the unity element $i(x)$ of R belongs to N and consequently $N = R$.

Hence M is a maximal ideal of R.

Ex. 3. *If R is a finite commutative ring (i.e., has only a finite number of elements) with unit element prove that every prime ideal of R is a maximal ideal of R.*

Solution. Let R be a finite commutative ring with unit element. Let S be a prime ideal of R. Then to prove that S is a maximal ideal of R.

Since S is a prime ideal of R, therefore the residue class ring R/S is an integral domain. Now

$$R/S = \{S + a : a \in R\}.$$

Since R is a finite ring, therefore R/S is a finite integral domain. But every finite integral domain is a field. Therefore R/S is a field. Since R is a commutative ring with unity and R/S is a field, therefore S is a maximal ideal of R

Ex. 4 *Give an example of a ring in which some prime ideal is not a maximal ideal.* **(I.A.S. 1970; Calicut 75)**

Solution. Let I $[x]$ be the ring of polynomials over the ring of integers I. Let S be the principal ideal of I $[x]$ generated by x *i.e.*, let $S=(x)$. We shall show that (x) is prime but not maximal. We have $S=(x)=\{x f(x) : f(x) \in \mathbf{I}[x]\}$.

First we shall prove that S is prime.

Let $a(x), b(x) \in \mathbf{I}[x]$ be such that $a(x)\, b(x) \in S$. Then there exists a polynomial $c(x) \in \mathbf{I}[x]$ such that

$$xc(x)=a(x)\, b(x). \qquad \ldots(1)$$

Let $a(x)=a_0+a_1x+a_2x^2+\ldots$, $b(x)=b_0+b_1x+b_2x^2+\ldots$

$c(x)=c_0+c_1x+c_2x^2+\ldots$. Then (1) becomes

$$x(c_0+c_1x+\ldots)=(a_0+a_1x+\ldots)(b_0+b_1x+\ldots).$$

Equating the constant term on both sides, we get

$$a_0b_0=0$$

$$\Rightarrow a_0=0 \text{ or } b_0=0 \qquad [\because \text{ I is without zero divisors}]$$

Now $a_0=0 \Rightarrow a(x)=a_1x+a_2x^2+\ldots$

$$\Rightarrow a(x)=x(a_1+a_2x+\ldots) \Rightarrow a(x) \in (x).$$

Similarly $b_0=0 \Rightarrow b(x)=b_1x+b_2x^2+\ldots$

$$\Rightarrow b(x)=x(b_1+b_2x+\ldots) \Rightarrow b(x) \in (x).$$

Thus $a(x)\, b(x) \in (x) \Rightarrow$ either $a(x) \in (x)$ or $b(x) \in (x)$. Hence (x) is a prime ideal.

Now we shall show that (x) is not a maximal ideal of I $[x]$. For this we must show an ideal N of I $[x]$ such that (x) is properly contained in N, while N itself is properly contained in I $[x]$. The ideal $N=(x, 2)$ serves this purpose.

Obviously $(x) \subseteq (x, 2)$. In order to show that (x) is properly contained in $(x, 2)$ we must show an element of $(x, 2)$ which is not in (x). Clearly $2 \in (x, 2)$. We shall show that $2 \notin (x)$. Let $2 \in (x)$. Then we can write,

$$2=xf(x) \text{ for some } f(x) \in \mathbf{I}[x].$$

Let $f(x)=a_0+a_1x+\ldots$

Then $2=xf(x) \Rightarrow 2=x(a_0+a_1x+\ldots)$

$$\Rightarrow 2=a_0x+a_1x^2+\ldots$$

$$\Rightarrow 2=0+a_0x+a_1x^2+\ldots$$

$$\Rightarrow 2=0 \qquad [\text{by equality of two polynomials}]$$

But $2 \neq 0$ in the ring of integers. Hence $2 \notin (x)$. Thus (x) is properly contained in $(x, 2)$.

Now obviously $(x, 2) \subseteq \mathbf{I}[x]$. In order to show that $(x, 2)$ is properly contained in $\mathbf{I}[x]$ we must show an element of $\mathbf{I}[x]$ which is not in $(x, 2)$. Clearly $1 \in \mathbf{I}[x]$. We shall show that $1 \notin (x, 2)$. Let $1 \in (x, 2)$. Then we have a relation of the form

$$1 = x\,f(x) + 2g(x), \qquad \text{where } f(x), g(x) \in \mathbf{I}[x].$$

Let $f(x) = a_0 + a_1x + \ldots,\ g(x) = b_0 + b_1x + \ldots$

Then $1 = x(a_0 + a_1x + \ldots) + 2(b_0 + b_1x + \ldots)$

$\Rightarrow 1 = 2b_0$ [Equating constant term on both sides]

But there is no integer b_0 such that $1 = 2b_0$.

Hence $1 \notin (x, 2)$. Thus $(x, 2)$ is properly contained in $\mathbf{I}[x]$.

Therefore (x) is not a maximal ideal of $\mathbf{I}[x]$.

Exercises

1. If U is an ideal of a ring R, then prove that R/U is a ring and is a homomorphic image of R. **(I.A.S. 1973)**
2. Let a be a given fixed real number and $\mathbf{R}$ be the ring of real numbers. Let ϕ be the mapping which associates with every polynomial $p(x)$—with real coefficients—the real number $p(a)$. Show that ϕ is a homomorphism from the ring $\mathbf{R}[x]$ onto the ring $\mathbf{R}$. **(I.A.S. 1975)**
3. Let R be the ring of all the real-valued continuous functions on the closed interval $[0, 1]$ of the real line, the compositions being the usual pointwise addition and multiplication of functions. Let M be a subset of R defined by :

   $$M = \{f \in R \mid f(\tfrac{1}{2}) = 0\}.$$

   Prove that M is a maximal ideal of R. **(I.A.S. 1969, 73)**
4. (a) Show that the set R of all elements of the form $m + in$, $m, n \in Z$, form a ring for addition and multiplication of complex numbers. Here Z is the ring of integers.

   (b) Is the map $\theta : R \to Z$, $\theta(m + in) = m$, a ring homomorphism ?

   (c) Give an example of a proper ideal in R. **(Poona 1973)**
5. Let I be an ideal in a ring R. If R is an integral domain, is R/I also an integral domain ? Justify. **(Poona 1973)**

§ 7. Euclidean Rings or Euclidean Domains.

Definition. *Let R be an integral domain i.e., let R be a commutative ring without zero divisors. Then R is said to be a Euclidean ring if to every non-zero element $a \in R$ we can assign a non-negative integer $d(a)$ such that :*

(*i*) *For all* $a, b \in R$, *both non-zero,* $d(ab) \geqslant d(a)$.

(*ii*) *For any* $a, b \in R$ *and* $b \neq 0$, *there exist* $q, r \in R$ *such that* $a = qb + r$ *where either* $r = 0$ *or* $d(r) < d(b)$.

(I.A S. 1973; Guru Nanak 82; Kanpur 80; Madras 83; Raj. 78; Andhra 75; Meerut 81, 82, 83, 84)

The second part of the above definition is known as **division algorithm**. Also we do not assign a value to $d(0)$. Thus $d(a)$ will remain undefined when $a = 0$. Also $d(a)$ will be called *d*-value of *a* and $d(a)$ must be some non-negative integer for every non-zero element $a \in R$

Example 1. *The ring of integers is a Euclidean ring.*

(Banaras 1971; Meerut 78, 81, 83, 84 P; Delhi 70)

Solution. Let $(\mathbf{I}, +, .)$ be the ring of integers where

$$\mathbf{I} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}.$$

Let the *d* function on the non-zero elements of **I** be defined as

$$d(a) = |a| \ \forall \ 0 \neq a \in \mathbf{I}.$$

Now if $0 \neq a \in \mathbf{I}$, then $|a|$ is a non-negative integer. Thus we have assigned a non-negative integer to every non-zero element $a \in \mathbf{I}$.

$[d(-5) = |-5| = 5,\ d(-1) = |-1| = 1,\ d(4) = |4| = 4 \text{ etc.}]$

Further if $a, b \in \mathbf{I}$ and are both non zero, then

$$|ab| = |a| \ |b|$$

$$\Rightarrow \ |ab| \geqslant |a| \qquad [\because \ |b| \geqslant 1 \text{ if } 0 \neq b \in \mathbf{I}]$$

$$\Rightarrow \ d(ab) \geqslant d(a).$$

Finally we know that if $a \in \mathbf{I}$ and $0 \neq b \in \mathbf{I}$, then there exist two integers *q* and *r* such that

$$a = qb + r \text{ where } 0 \leqslant r < |b|$$

i.e., where either $r = 0$ or $1 \leqslant r < |b|$

i.e., where either $r = 0$ or $d(r) < d(b)$.

It should be noted that $d(b) = |b|$ and if *r* is a positive integer then $r = |r| = d(r)$.

Therefore the ring of integers is a Euclidean ring.

Example 2. *The ring of polynomials over a field is a Euclidean ring.*

(Kanpur 1980; Madras 83; Meerut 79, 83 P)

Solution. Let $F[x]$ be the ring of polynomials over a field *F*. Let the *d* function on the non-zero polynomials in $F[x]$ be defined as

$$d[f(x)] = \deg f(x), \ \forall \ 0 \neq f(x) \in F[x].$$

Now if $0 \neq f(x) \in F[x]$, then $\deg f(x)$ is a non-negative integer.

Thus we have assigned a non-negative integer to every non-zero element $f(x)$ in $F[x]$.

Further if $f(x), g(x) \in F[x]$ and are both non-zero polynomials, then

$$\deg [f(x)\, g(x)] = \deg f(x) + \deg g(x)$$

$$\Rightarrow \deg [f(x)\, g(x)] \geqslant \deg f(x) \qquad [\because \ \deg g(x) \geqslant 0]$$

$$\Rightarrow d\,[f(x)\, g(x)] \geqslant d\,[f(x)].$$

Finally we know that if $f(x) \in F[x]$ and $0 \neq g(x) \in F[x]$, then there exist two polynomials $q(x)$ and $r(x)$ in $F[x]$ such that

$$f(x) = q(x)\, g(x) + r(x)$$

where either $r(x)=0$ or $\deg r(x) < \deg g(x)$

i.e., where either $r(x)=0$ or $d\,[r(x)] < d\,[g(x)]$.

Hence the ring of polynomials over a field is a Euclidean ring.

Example 3. *Every field is a Euclidean ring.*

Solution. Let F be any field. Let the d function on the non-zero elements of F be defined as

$$d(a)=0 \ \forall \ 0 \neq a \in F.$$

Thus we have assigned the integer zero to every non-zero element in F.

If a and b are two non-zero elements in F then ab is also a non-zero element in F. We have therefore

$$d(ab) = 0 = d(a).$$

Thus we have $d(ab) \geqslant d(a)$.

Finally if $a \in F$ and $0 \neq b \in F$, then we can write

$$a = (ab^{-1})\, b + 0$$

i.e., $$a = qb + r \text{ where } q = ab^{-1} \text{ and } r = 0.$$

Hence every field is a Euclidean ring.

Example 4. *The ring of Gaussian integers is a Euclidean ring.*

(I.A.S. 1973; Gujrat 70; Mysore 70; Delhi 70; Kanpur 80; Meerut 79, 80, 81, 82 P)

Solution. Let $(G, +, .)$ be the ring of Gaussian integers where

$$G = \{x + iy : x, y \in I\}$$

Let the d function on the non-zero elements of G be defined as

$$d(x+iy) = x^2 + y^2 \ \forall \ 0 + i0 \neq x + iy \in G.$$

Now if $x+iy$ is a non-zero element of G, then (x^2+y^2) is a non-negative integer. Thus we have assigned a non-negative integer to every non-zero element of G.

If $x+iy$ and $m+in$ are two non-zero elements of G, then

$$d\,[(x+iy)\,(m+in)] = d\,[(xm-ny) + i\,(my+xn)]$$

$$= (xm-ny)^2+(my+xn)^2 = x^2m^2+n^2y^2+m^2y^2+x^2n^2$$
$$= (x^2+y^2)(m^2+n^2)$$
$$\geqslant x^2+y^2. \qquad [\because \; m^2+n^2 \geqslant 1]$$

Thus $d[(x+iy)(m+in)] \geqslant d(x+iy)$.

Now to show the existence of division algorithm in G.

Let $\alpha \in G$ and let β be a non-zero element of G. Let $\alpha = x+iy$ and $\beta = m+in$. Define a complex number λ by the equation

$$\lambda = \frac{\alpha}{\beta} = \frac{x+iy}{m+in} = \frac{(x+iy)(m-in)}{m^2+n^2} = p+iq,$$

where p, q are rational numbers.

Here λ is not necessarily a Gaussian integer.

Also division by β is possible since $\beta \neq 0$.

Let p' and q' be the nearest integers to p and q respectively. Then obviously $|p-p'| \leqslant \frac{1}{2}$, $|q-q'| \leqslant \frac{1}{2}$.

Let $\lambda' = p'+iq'$. Then λ' is a Gaussian integer.

Now $\lambda = \dfrac{\alpha}{\beta} \Rightarrow \alpha = \lambda\beta$

$$\Rightarrow \alpha = \lambda'\beta + \lambda\beta - \lambda'\beta.$$

Thus $\alpha = \lambda'\beta + (\lambda - \lambda')\beta$. ...(1)

Since α, β, λ' are Gaussian integers, therefore from (1) it implies that $(\lambda-\lambda')\beta$ is also a Gaussian integer.

Now if p and q are integers then $p=p'$, $q=q'$.

So $\lambda-\lambda' = (p-p')+i(q-q') = 0+i0$. Thus $(\lambda-\lambda')\beta = 0+i0$.

If p and q are not both integers, then $(\lambda-\lambda')\beta$ is a non-zero Gaussian integer and we have

$$d[(\lambda-\lambda')\beta] = d[\{(p-p')+i(q-q')\}(m+in)]$$
$$= [(p-p')^2+(q-q')^2](m^2+n^2) = [(p-p')^2+(q-q')^2]\, d(\beta)$$
$$\leqslant [\tfrac{1}{4}+\tfrac{1}{4}]\, d(\beta) \qquad [\because (p-p')^2 \leqslant \tfrac{1}{4},\ (q-q')^2 \leqslant \tfrac{1}{4}]$$
$$= \tfrac{1}{2} d(\beta) < d(\beta).$$

Thus $\alpha = \lambda'\beta + (\lambda-\lambda')\beta$ where λ' and $(\lambda-\lambda')\beta$ are Gaussian integers and either $(\lambda-\lambda')\beta = 0$

or $\qquad d[(\lambda-\lambda')\beta] < d(\beta)$.

Hence the ring of Gaussian integers is a Euclidean ring.

§ 8. Properties of Euclidean rings.

Theorem 1. *Every Euclidean ring is a principal ideal ring.*

(I.A.S. 1972; Guru Nanak 82; Meerut 79, 82 P, 84; Madurai 78; Sambalpur 77; Andhra 75; Raj. 77; Gujrat 70)

Proof. Let R be a Euclidean ring. Let S be an arbitrary ideal of R. If S is the null ideal, then $S=(0)$ *i.e.*, the ideal of R generated by 0. Therefore S is a principal ideal. So let us suppose that S is not a null ideal. Then there exist elements in S not equal to zero. Let b be a non-zero element in S such that $d(b)$ is minimum *i e.*, there exists no element c in S such that $d(c) < d(b)$. We shall show that $S=(b)$ *i.e.*, S is nothing but the ideal generated by b.

Let a be any element of S. Then by definition of Euclidean ring there exist elements q and r in R such that

$$a=qb+r \text{ where either } r=0 \text{ or } d(r) < d(b).$$

Now $q \in R, b \in S \Rightarrow qb \in S$ because S is an ideal.

Further $a \in S, qb \in S \Rightarrow a-qb=r \in S$.

Thus $r \in S$ and we have either $r=0$ or $d(r) < d(b)$.

If $r \neq 0$, then $d(r) < d(b)$ which contradicts our assumption that no element in S has d-value smaller than $d(b)$. Therefore we must have $r=0$.

Then $$a=qb.$$

Thus every element a in S is a multiple of the generating element b. Thus $a \in S \Rightarrow a \in (b)$. Therefore $S \subseteq (b)$.

Again if xb is any element of (b), then $x \in R$.

Now $x \in R, b \in S \Rightarrow xb \in S$ Therefore $(b) \subseteq S$.

Hence $$S=(b)$$

Thus every ideal S in R is a principal ideal. Therefore R is a principal ideal ring.

Theorem 2. *Every Euclidean ring possesses unity element.*

(Marathwada 1972; Meerut 73, 80, 81; Banaras 68)

Proof. Let R be a Euclidean ring. Obviously R is an ideal of R. Therefore there exists an element $u_0 \in R$ such that $R=(u_0)$ *i.e.*, there exists an element $u_0 \in R$ such that every element in R is a multiple of u_0. Since, in particular, $u_0 \in R$ therefore there exists an element $c \in R$ such that $u_0=u_0c$. We shall show that c is the required unity element. Let now a be any element of R. Since $a \in R$, therefore there exists some $x \in R$ such that $a=u_0x$.

$$\begin{aligned} \text{Now} \quad ac &= (u_0x)\, c && [\because \ a=u_0x] \\ &= (u_0c)\, x && [\because \ R \text{ is a commutative ring}] \\ &= u_0x && [\because \ u_0=u_0c] \\ &= a && [\because \ a=u_0x]. \end{aligned}$$

Thus we have $ac=a=ca \ \forall \ a \in R$.

Hence c is the unity element.

Theorem 3. *Let R be a Euclidean ring and a and b be any two elements in R, not both of which are zero. Then a and b have a greatest common divisor d which can be expressed in the form*

$$d=\lambda a+\mu b \text{ for some } \lambda, \mu \in R.$$

(Kanpur 1980; Meerut 74, 76; Madras 83)

Proof. Consider the set

$$S=\{sa+tb : s, t \in R\}.$$

We claim that S is an ideal of R. The proof is as follows :

Let $x=s_1a+t_1b$, and $y=s_2a+t_2b$ be any two elements of S.

Then $s_1, t_1, s_2, t_2 \in R$. We have

$x-y=(s_1a+t_1b)-(s_2a+t_2b)=(s_1-s_2)\ a+(t_1-t_2)\ b \in S$

since s_1-s_2 and t_1-t_2 are both elements of R.

Thus S is a subgroup of R with respect to addition.

Also if u be any element of R, then

$xu=ux=u\ (s_1a+t_1b)=(us_1)\ a+(ut_1)\ b \in S$ since $us_1, ut_1 \in R$.

Therefore S is an ideal of R. Now every ideal in R is a principal ideal. Therefore there exists an element d in S such that every element in S is a multiple of d.

Since $d \in S$, therefore from (1), we see that there exist elements $\lambda, \mu \in R$ such that $d=\lambda a+\mu b$.

Now R is a ring with unity element 1.

$\therefore$ Putting $s=1$, $t=0$ in (1), we see that $a \in S$. Also putting $s=0$, $t=1$ in (1), we see that $b \in S$.

Now a, b are elements of S. Therefore they are both multiples of d. Hence $d \mid a$ and $d \mid b$.

Now suppose $c \mid a$ and $c \mid b$.

Then $c \mid \lambda a$ and $c \mid \mu b$. Therefore c is also a divisor of $\lambda a+\mu b$ *i.e.*, c is a divisor of d.

Thus d is a greatest common divisor of a and b.

Theorem 4. *Let a, b and c be any elements of a Euclidean ring R. Let $(a, b)=1$ i.e., let the greatest common divisor of a and b be 1. If $a \mid bc$, then $a \mid c$.* **(Kumayun 1977; Madras 83)**

Proof. If the greatest common divisor of a and b is 1, then by theorem 3 there exist elements λ and μ in R such that

$$1=\lambda a+\mu b. \quad \text{...(1)}$$

Multiplying both members of (1) by c, we get

$$c=\lambda ac+\mu bc. \quad \text{...(2)}$$

But $a \mid bc$, so there exists an element $q \in R$ such that

$$bc = qa.$$

Substituting this value of bc in (2), we get

$$c = \lambda ac + \mu qa = (\lambda c + \mu q)\, a,$$

which shows that a is a divisor of c. Hence the theorem.

Theorem 5. *If p is a prime element in the Euclidean ring R and $p \mid ab$ where $a, b \in R$ then p divides at least one of a or b.*

Proof. If p divides a, we are nothing to prove. So suppose that p does not divide a. Since p is prime and p does not divide a, therefore p and a are relatively prime *i.e.*, the greatest common divisor of p and a is 1. Hence by theorem 4, we get that $p \mid b$.

Corollary. *If p is a prime element in the Euclidean ring R and p divides the product $a_1 a_2 \ldots a_n$ of elements in R, then p divides at least one of $a_1, a_2, \ldots, a_n$.*

The result follows immediately by repeated application of theorem 5.

Theorem 6. *Let R be a Euclidean ring. Let a and b be two non-zero elements in R. Then*

(i) *if b is a unit in R, $d(ab) = d(a)$.*

(ii) *if b is not a unit in R, $d(ab) > d(a)$.*

(Meerut 1982, 83)

Proof. (i) By the definition of Euclidean ring, we have

$$d(ab) \geqslant d(a). \qquad \ldots(1)$$

Now suppose that b is a unit in R. Then b is inversible and b^{-1} exists. We can write

$$a = (ab)\, b^{-1}.$$

$$\therefore \quad d(a) = d\,[(ab)\, b^{-1}]$$

But by the definition of Euclidean ring, we have

$$d\,[(ab)\, b^{-1}] \geqslant d(ab).$$

$$\therefore \quad d(a) \geqslant d(ab). \qquad \ldots(2)$$

From (1) and (2), we conclude that

$$d(ab) = d(a).$$

(ii) Suppose now that b is not a unit in R. Since a and b are non-zero elements of the Euclidean ring R, therefore ab is also a non-zero element of R. Now $a \in R$ and $0 \neq ab \in R$, therefore by definition of Euclidean ring there exist elements q and r in R such that

$$a = q\,(ab) + r \qquad \ldots(3)$$

where either $r = 0$ or $d(r) < d(ab)$.

If $r=0$, then

$$a=qab$$
$$\Rightarrow a-qab=0 \Rightarrow a\,(1-qb)=0$$
$$\Rightarrow 1-qb=0 \quad [\because \; a\neq 0 \text{ and } R \text{ is free of zero divisors}]$$
$$\Rightarrow qb=1 \Rightarrow b \text{ is inversible} \Rightarrow b \text{ is a unit in } R.$$

Thus we get a contradiction. Hence r cannot be zero. Therefore we must have

$$d(r) < d(ab)$$

i.e.,
$$d(ab) > d(r). \qquad \ldots(4)$$

Also from (3), we have $r=a-qab=a\,(1-qb)$.

$$\therefore \quad d(r)=d[a\,(1-qb)].$$

But
$$d[a\,(1-qb)] \geqslant d\,(a).$$

$$\therefore \quad d\,(r) \geqslant d(a). \qquad \ldots(5)$$

From (4) and (5), we conclude that $d\,(ab)>d(a)$.

Theorem 7. *The necessary and sufficient condition that the non-zero element a in the Euclidean ring R is a unit is that*

$$d\,(a)=d\,(1).$$

Proof. Let a be a unit in R. Then to prove that $d\,(a)=d(1)$. By the definition of Euclidean ring

$$d\,(1a) \geqslant d(1)$$
$$\Rightarrow d\,(a) \geqslant d(1). \qquad \ldots(1)$$

Since a is a unit in R, therefore a^{-1} exists and we have

$$1=aa^{-1}$$
$$\Rightarrow d\,(1)=d\,(aa^{-1}).$$

But
$$d\,(aa^{-1}) \geqslant d(a).$$

$$\therefore \quad d\,(1) \geqslant d(a). \qquad \ldots(2)$$

From (1) and (2), we conclude that $d\,(a)=d(1)$.

Conversely let $d\,(a)=d\,(1)$. Then to prove that a is a unit in R. If a is not a unit in R, then by theorem 6, we have

$$d\,(1a) > d\,(1)$$
$$\Rightarrow d\,(a) > d(1).$$

Thus we get a contradiction. Hence a must be a unit in R.

Theorem 8. *Let R be a Euclidean ring. Then every non-zero element in R is either a unit in R or can be written as a product of a finite number of prime elements of R.*

(I.A.S. 1971; Meerut 73)

Proof. Let a be a non-zero element of R. We are to prove that either a is a unit in R or it can be written as a product of a finite number of prime elements of R. We shall prove the result by induction on $d(a)$ *i.e.*, by induction on the d-value of a.

Let us first start the induction. We have $a=1a$. Therefore $d(a) \geqslant d(1)$. Thus 1 is an element in R which has the minimal d-value. If $d(a)=d(1)$, then a is a unit in R [See theorem 7]. Thus the result of the theorem is true if $d(a)=d(1)$ and so we have started the induction.

Now assume as our induction hypothesis that the theorem is true for all non-zero elements $x \in R$ such that $d(x) < d(a)$. Then we shall show that the theorem is true for a also. If a is a prime element of R, the theorem is obviously true. So suppose that a is not prime Then we can write $a=bc$ where neither b nor c is a unit in R. Since both b and c are not units in R, therefore $d(bc)>d(b)$ and $d(bc) > d(c)$. But $d(a)=d(bc)$. Therefore we have $d(b) < d(a)$ and $d(c)<d(a)$. So by our induction hypothesis each of b and c can be written as a product of a finite number of prime elements of R. Let $b=p_1p_2 \ldots p_n$, $c=q_1q_2 \ldots q_m$ where the p's and q's are prime elements of R. Then $a=bc=p_1p_2 \ldots p_n\, q_1q_2 \ldots q_m$.

Thus we have written a as a product of a finite number of prime elements of R. This completes the induction and so the theorem has been proved.

Theorem 9. Unique Factorization theorem. *Let R be a Euclidean ring and a be a non-zero non-unit element in R. Suppose that*

$$a=p_1p_2 \ldots p_m=q_1q_2 \ldots q_n$$

where the p's and q's are prime elements of R. Then $m=n$ and each p is an associate of some q and each q is an associate of some p.

Proof. We have $p_1p_2 \ldots p_m=q_1q_2 \ldots q_n$. Now p_1 is a divisor of $p_1p_2 \ldots p_m$. Therefore p_1 is also a divisor of $q_1q_2 \ldots q_n$. By Cor. to Theorem 5, p_1 must divide at least one of $q_1, q_2, \ldots, q_n$. Since R is a commutative ring, therefore without loss of generality we may suppose that p_1 divides q_1. But p_1 and q_1 are both prime elements in R. Therefore p_1 and q_1 must be associates and we have $q_1=up_1$ where u is a unit in R. Thus we have

$$p_1p_2 \cdots p_m=up_1q_2 \cdots q_n.$$

Cancelling $0 \neq p_1$ from both sides, we get

$$p_2p_3 \cdots p_m=uq_2q_3 \cdots q_n. \qquad \ldots(1)$$

Now we can repeat the above argument on the relation (1) with p_2. If $n > m$, then after m steps the left hand side becomes 1 while the right hand side reduces to a product of some units in

R and certain number of q's (the excess of n over m). But the q's are prime elements of R and so they are not units in R. So the product of some units in R and certain number of q's cannot be equal to 1. Therefore n cannot be greater than m.

Thus $n \leqslant m$.

Similarly interchanging the roles of p's and q's we get

$m \leqslant n$. Hence $m = n$.

Also in the above process we have shown that every p is an associate of some q and conversely every q is an associate of some p. Hence the theorem has been completely established.

Note. Combining theorems 8 and 9, we can say that *every non-zero element in a Euclidean ring R can be uniquely written (upto associates) as a product of prime elements or is a unit in R. Therefore a Euclidean ring is a Unique Factorization Domain.*

(Dibrugarh 1967; Meerut 1973)

Theorem 10. *An ideal S of the Euclidean ring R is maximal iff S is generated by some prime element of R.*

(I A.S. 1972; Meerut 70, 72; Calicut 75)

Proof. We know that every ideal of a Euclidean ring R is a principal ideal. Suppose S is an ideal of R generated by p so that $S = (p)$. Now we are to prove that

(i) S is maximal if p is a prime element of R.

(ii) p is prime if S is maximal.

First we shall prove (i). Let p be a prime element of R such that $(p) = S$. Let T be an ideal of R such that $S \subseteq T \subseteq R$. Since T is also a principal ideal of R, so let $T = (q)$ where $q \in R$.

Now $\quad S \subseteq T \Rightarrow (p) \subseteq (q) \Rightarrow p \in (q)$

$\Rightarrow p = xq$ for some $x \in R \Rightarrow q \mid p$.

Since p is prime, therefore either q should be a unit in R or q should be an associate of p.

If q is a unit in R, then $T = (q) = R$.

If q is an associate of p, then $T = (q) = (p) = S$.

Thus either $T = R$ or $T = S$.

Now we shall prove (ii). Let $(p) = S$ be a maximal ideal. We are to show that p is prime. Let us suppose that p is composite *i.e.*, p is not prime.

Let $p = mn$ where neither m nor n is a unit in R.

Now $p = mn \Rightarrow m \mid p$

$\Rightarrow (p) \subseteq (m)$.

But $(m) \subsetneq R$. Therefore we have $(p) \subseteq (m) \subsetneq R$.

But (p) is a maximal ideal, therefore we should have either

$$(m) = (p) \quad \text{or} \quad (m) = R.$$

If $R = (m)$, then $R \subseteq (m)$.

$\therefore \quad 1 \in R \Rightarrow 1 \in (m)$

$\Rightarrow 1 = ym$ for some $y \in R$

$\Rightarrow m$ is inversible $\Rightarrow m$ is a unit in R.

Thus we get a contradiction.

If $(m) = (p)$, then $m \in (p)$. Therefore $m = lp$ for some $l \in R$.

$\therefore \quad p = mn = lpn = pln$.

$\therefore \quad p(1 - ln) = 0$

$\Rightarrow 1 - ln = 0$ $\quad$ [$\because$ $p \neq 0$ and R is without zero divisors]

$\Rightarrow ln = 1 \Rightarrow n$ is inversible $\Rightarrow n$ is a unit in R.

This is again a contradiction. Hence p must be a prime element of R.

§ 9. Polynomial rings over Unique Factorization Domains.

We have already defined a unique factorization domain. For the sake of convenience we repeat the definition here.

Unique Factorization Domain. Definition. *An integral domain R, with unity element 1 is a unique factorization domain if*

(*a*) *any non-zero element in R is either a unit or can be written as the product of a finite number of irreducible* (*prime*) *elements of R.*

(*b*) *the decomposition in part* (*a*) *is unique upto the order and associates of the irreducible elements.*

(Punjab 1968 ; Meerut 70 ; Calicut 75)

In general commutative rings we have defined the greatest common divisors of elements. But the difficulty is that in an arbitrary commutative ring these might not exist. However, in unique factorization domain their existence is assured. Further we know that in an integral domain with unity in case a greatest common divisor of some elements exists, it is unique apart from the distinction between associates.

Theorem 1. *Let R be a unique factorization domain and a and b be any two elements in R, not both of which are zero. Then a and b have a greatest common divisor (a, b) in R. Moreover, if a and b are relatively prime (i.e., (a, b) is a unit in R), then $a \mid bc \Rightarrow a \mid c$.*

Proof. Suppose a and b are any two elements, not both of which are zero, of a unique factorization domain R. If one of a and b, say, b is 0, then obviously a is the greatest common divisor of a and b. If any of a and b, say a, is a unit in R, then obviously a is the greatest common divisor of a and b. So let us suppose that neither $a=0$ nor $b=0$ and none of these is a unit in R. Then each of a and b can be uniquely expressed as the product of a finite number of irreducible elements of R. Let

$$a=p_1^{m_1}p_2^{m_2}\dots p_r^{m_r}, \qquad \dots(1)$$

and $$b=p_1^{n_1}p_2^{n_2}\dots p_r^{n_r}, \qquad \dots(2)$$

where we have arranged the expressions in such a way that the same irreducible factors $p_1, p_2, \dots, p_r$ appear in both. Note that we can definitely do so because the integer 0 can be used as power in any case, if necessary. The elements $p_1, p_2, \dots, p_r$ are all different primes and $m_1, m_2, \dots, m_r, n_1, n_2, \dots, n_r$ are all integers $\geqslant 0$.

Let $g_i=\text{minimum } (m_i, n_i)$, where $i=1, 2, \dots, r$. Then obviously

$$p_1^{g_1}p_2^{g_2} \quad p_r^{g_r}$$

is the greatest common divisor of a and b.

This proves the existence of greatest common divisor.

Now suppose that a and b are relatively prime *i.e.*, the greatest common divisor of a and b is a unit in R. Also suppose that

$$a \mid bc.$$

If a is a unit in R, then obviously a is a divisor of c. So let a be not a unit in R. Then a can be uniquely expressed as the product of a finite number of prime elements of R. Let

$$a=q_1 q_2 \cdot q_s,$$

where $q_1, q_2, \dots, q_s$ are prime elements of R.

We have $$a \mid bc \Rightarrow bc=ka \text{ for some } k \in R$$
$$\Rightarrow bc=kq_1 q_2 \dots q_s. \qquad \dots(3)$$

Since each element of R can be uniquely expressed as the product of a finite number of prime elements of R, therefore each of the prime elements $q_1, q_2, \dots, q_s$ must occur as a factor of either b or c. But none of $q_1, q_2, \dots, q_s$ can be a factor of b because otherwise a and b will not remain relatively prime. Therefore each of $q_1, q_2, \dots, q_s$ must be a factor of c. Hence

$$q_1 q_2 \dots q_s \text{ is a divisor of } c \Rightarrow a \mid c.$$

Note. In a similar manner we can prove that if $a_1, \dots, a_n$ are any n elements of a unique factorization domain, they possess a greatest common divisor which will be unique apart from the distinction between associates. Thus if g_1, g_2 are two greatest common divisors of these n elements, then by the definition of greatest common divisor, we have

$$g_1 \mid g_2 \text{ and } g_2 \mid g_1$$
$$\Rightarrow g_1 \text{ and } g_2 \text{ are associates}$$
$$\Rightarrow g_1 = ug_2 \text{ where } u \text{ is a unit in } R.$$

Thus the greatest common divisor of some elements is unique within units of R.

Theorem 2. *If a is a prime element of a unique factorization domain R and b, c are any elements of R, then*

$$a \mid bc \Rightarrow a \mid b \text{ or } a \mid c.$$

Proof. If $a \mid b$, then obviously the theorem is proved. So let a be not a divisor of b. Since a is a prime element of R and a is not a divisor of b, therefore we claim that a and b are relatively prime. Since a is a prime element of R, therefore the only divisors of a are the associates of a or the units of R. Now an associate of a cannot be a divisor of b otherwise a itself will be a divisor of b while we have assumed that a is not a divisor of b. Thus the units of R are the only divisors of a which also divide b. Therefore the greatest common divisor of a and b is a unit of R.

Since a and b are relatively prime, therefore by theorem 1, we have

$$a \mid bc \Rightarrow a \mid c.$$

This completes the proof of the theorem.

Polynomial rings over unique factorization domains. Let R be a unique factorization domain. Since R is an integral domain with unity, therefore $R[x]$ is also an integral domain with unity. Also any unit, (inversible element) in $R[x]$ must already be a unit in R. Thus the only units in $R[x]$ are the units of R. A polynomial $p(x) \in R[x]$ is *irreducible* over R *i e*, irreducible as an element of $R[x]$ if whenever $p(x)=a(x)\, b(x)$ with $a(x), b(x) \in R[x]$, then one of $a(x)$ or $b(x)$ is a unit in $R[x]$ *i e.*, a unit in R. For example, if **I** is the ring of integers, then **I** is a unique factorization domain. The polynomial $2x^2+4 \in \mathbf{I}[x]$ is a reducible element of **I** $[x]$. We have $2x^2+4=2(x^2+2)$. Neither 2 nor x^2+2 is a unit in **I** $[x]$. On the other hand the polynomial $x^2+1 \in \mathbf{I}[x]$ is an irreducible element of **I** $[x]$.

Content of a polynomial.

Definition *Let $f(x)=a_0+a_1x+a_2x^2+\dots+a_nx^n$ be a polynomial* *over a unique factorization domain R. Then the content of $f(x)$ denoted by $c(f)$, is defined as the greatest common divisor of the coefficients $a_0, a_1, \ldots, a_n$ of $f(x)$.* Obviously the content of $f(x)$ is unique within units of R. Thus if c_1 and c_2 are two contents of $f(x)$, then we must have $c_1 = uc_2$ where u is some unit in R.

Primitive polynomial. Definition. *Let R be a unique factorization domain. Then a polynomial $f(x) = a_0 + a_1x + \ldots + a_nx^n \in R[x]$ is said to be primitive if the greatest common divisor of its coefficients $a_0, a_1, \ldots, a_n$ is a unit in R.* Thus a polynomial $f(x)$ is primitive if its content is 1 (that is a unit in R). If

$$f(x) = x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n$$

is a monic polynomial over R, then obviously $f(x)$ is primitive.

If **I** is the ring of integers, then $3x^3 - 5x^2 + 7$ is a primitive member of **I** $[x]$ while $2x^2 - 4x + 8$ is not a primitive member of **I**$[x]$.

Every irreducible polynomial of positive degree belonging to $R[x]$ is necessarily primitive. But an irreducible polynomial of zero degree may not be primitive. For example $3 \in$ **I** $[x]$ is irreducible but it is not primitive. Further a primitive polynomial may not be irreducible. For example $x^2 + 5x + 6 \in$ **I** $[x]$ is primitive and it is not irreducible. We have $x^2 + 5x + 6 = (x+2)(x+3)$.

Theorem 3. *Let R be a unique factorization domain. Prove that every non-zero member $f(x)$ of $R[x]$ can be written as $f(x) = gf_1(x)$ where $g = c(f)$ and where $f_1(x)$ is primitive. Also prove that this decomposition of $f(x)$ as an element of R by a primitive polynomial in $R[x]$ is unique apart from the distinction between associates.*

Proof. Let R be a unique factorization domain and let

$$f(x) = a_0 + a_1x + \ldots + a_nx^n \in R[x].$$

Since R is a unique factorization domain, therefore the elements $a_0, a_1, \ldots, a_n \in R$ must possess a greatest common divisor. Let $g \in R$ be the greatest common divisor of these elements. Then $g = c(f)$. Let

$$a_i = gb_i \text{ where } i = 0, 1, \ldots, n.$$

Then $f(x) = gb_0 + gb_1x + \ldots + gb_nx^n = g\,[b_0 + b_1x + \ldots + b_nx^n]$.

Since g is the greatest common divisor of $a_0, a_1, \ldots, a_n$, therefore the elements $b_0, b_1, \ldots, b_n$ can have no common factor other than units of R. Consequently the polynomial

$$f_1(x) = b_0 + b_1x + \ldots + b_nx^n$$

is a primitive member of $R[x]$. Thus we have $f(x) = gf_1(x)$, where

$g \in R$ and $f_1(x) \in R[x]$ is primitive.

Now we come to the uniqueness part of the theorem.

If possible, let

$$f(x)=h\,f_2(x) \text{ where } h \in R \text{ and } f_2(x) \in R[x]$$

is primitive.

Then $g f_1(x)=h f_2(x)$. ...(1)

Since $f_1(x)$ and $f_2(x)$ are both primitive, therefore the content of the polynomial on the left hand side of (1) is g and the content of the polynomial on the right hand side of (1) is h. But the content of a polynomial is unique upto associates. Therefore g and h are associates

$\Rightarrow$ $g=hu$ where u is some unit in R

$\Rightarrow$ $hu\,f_1(x)=hf_2(x)$

$\Rightarrow$ $uf_1(x)=f_2(x)$

[by left cancellation law in the integral domain $R[x]$, since $h \neq 0$]

$\Rightarrow$ $f_1(x)$ and $f_2(x)$ are associates.

Hence the theorem.

Theorem 4. *If R is a unique factorization domain, then the product of two primitive polynomials in $R[x]$ is again a primitive polynomial in $R[x]$.* **(Madras 1983, Guru Nanak 82)**

Proof. Let

$$f(x)=a_0+a_1x+\ldots+a_nx^n$$

and

$$g(x)=b_0+b_1x+\ldots+b_mx^m$$

be two primitive polynomials in $R[x]$.

Let $h(x)=f(x)\,g(x)=c_0+c_1x+\ldots+c_{m+n}x^{m+n}$.

Suppose $h(x)$ is not primitive. Then all the coefficients of $h(x)$ must be divisible by some prime element p of R. Since $f(x)$ is primitive, therefore the prime element p must not divide some coefficient of $f(x)$. Let a_i be the first coefficient of $f(x)$ which p does not divide. Similarly let b_j be the first coefficient of $g(x)$ which p does not divide. In $f(x)\,g(x)$, the coefficient of x^{i+j} is

$$c_{i+j}=a_ib_j+(a_{i-1}\,b_{j+1}+a_{i-2}\,b_{j+2}+\ldots+a_0b_{i+j}) + (a_{i+1}\,b_{j-1}+a_{i+2}\,b_{j-2}+\ldots+a_{i+j}\,b_0).$$

From this relation, we get

$$a_ib_j=c_{i+j}-(a_{i-1}\,b_{j+1}+a_{i-2}\,b_{j+2}+\ldots+a_0b_{i+j}) - (a_{i+1}\,b_{j-1}+a_{i+2}\,b_{j-2}+\ldots+a_{i+j}\,b_0) \quad \ldots(1)$$

Now by our choice of a_i, p is a divisor of each of the elements $a_0, a_1, \ldots, a_{i-1}$. Therefore $p \mid (a_{i-1}\,b_{j+1}+a_{i-2}\,b_{j+2}+\ldots+a_0\,b_{i+j})$.

Similarly, by our choice of b_j, p is a divisor of each of the elements $b_0, b_1, \dots, b_{j-1}$. Therefore $p \mid (a_{i+1} b_{j-1} + a_{i+2} b_{j-2} + \dots + a_{i+j} b_0)$.

Also by assumption $p \mid c_{i+j}$.

Hence from (1), we get

$$p \mid a_i b_j$$

$$\Rightarrow p \mid a_i \text{ or } p \mid b_j, \text{ since } p \text{ is a prime element of } R.$$

But this is nonsense because according to our assumption p is not a divisor of a_i and also p is not a divisor of b_j.

Hence $h(x)$ must be primitive. This proves the theorem.

Theorem 5 *If R is a unique factorization domain and if $f(x)$, $g(x)$ are in $R[x]$, then*

$$c(fg) = c(f)\, c(g) \text{ (upto units).}$$

Proof. The polynomial $f(x)$ in $R[x]$ can be written as $f(x) = af_1(x)$, where $a = c(f)$ and $f_1(x)$ is primitive. Similarly the polynomial $g(x)$ can be written as $g(x) = bg_1(x)$, where $b = c(g)$ and $g_1(x)$ is primitive. Then

$$f(x)\, g(x) = ab\, f_1(x)\, g_1(x) \qquad \dots(1)$$

Since $f_1(x)$ and $g_1(x)$ are both primitive, therefore $f_1(x)\, g_1(x)$ is also primitive. [Refer theorem 4.]

Therefore from (1), we see that the content of $f(x)\, g(x)$ is either ab or some associate of ab. Thus the content of $f(x)\, g(x)$ is ab (upto units). Therefore

$$c(fg) = ab = c(f)\, c(g).$$

This proves the theorem.

Field of Quotients of a unique factorization domain. If R is a unique factorization domain, then R is necessarily an integral domain. Therefore R has a field of quotients. Throughout this section we shall denote the field of quotients of R by F. We can consider $R[x]$ to be a subring of $F[x]$.

Theorem 6. *If R is an integral domain (not necessarily a unique factorization domain) and F is its field of quotients, then any element $f(x)$ in $F[x]$ can be written as*

$$f(x) = \frac{f_0(x)}{a},$$

where $f_0(x) \in R[x]$ and where $a \in R$.

Proof. Let F be the field of quotients of an integral domain R. Then

$$= \left\{ \frac{p}{q} : p \in R,\ 0 \neq q \in R \right\}.$$

Let $f(x)$ be an element of $F[x]$. Let

$f(x)=\frac{a_0}{b_0}+\frac{a_1}{b_1}x+\ldots+\frac{a_n}{b_n}x^n$, where $a_0, a_1, \ldots, a_n \in R$ and $b_0, b_1, \ldots, b_n$ are non-zero elements of R.

Now $b_0, b_1, \ldots, b_n$ are also non-zero elements of F. So each of them must be inversible. Further $b_0 b_1 \ldots b_n$ is also a non-zero element of F and so it is also inversible. Then we can write

$$f(x)=\frac{b_0 b_1 \ldots b_n}{b_0 b_1 \ldots b_n}\left(\frac{a_0}{b_0}+\frac{a_1}{b_1}x+\ldots+\frac{a_n}{b_n}x^n\right)$$

$$=\frac{(a_0 b_1 b_2 \ldots b_n)+(b_0 a_1 b_2 \ldots b_n)x+\ldots+(b_0 b_1 \ldots b_{n-1}a_n)x^n}{b_0 b_1 \ldots b_n}$$

$$=\frac{f_0(x)}{a}$$

where obviously $f_0(x)=(a_0 b_1 b_2 \ldots b_n)+(b_0 a_1 b_2 \ldots b_n)x+\ldots$
$+(b_0 b_1 \ldots b_{n-1}a_n)x^n$ is in $R[x]$ and $a=b_0 b_1 \ldots b_n$ is in R.

Theorem 7. (Gauss' Lemma). *Let F be the field of quotients of a unique factorization domain R. If the primitive polynomial $f(x) \in R[x]$ can be factored as the product of two polynomials having coefficients in F, then it can be factored as the product of two polynomials having coefficients in R.*

Proof. Let R be a unique factorization domain and F be its field of quotients. Let $f(x) \in R[x]$ be primitive.

Let $f(x)=g(x)\,h(x)$ where $g(x)$ and $h(x)$ have coefficients in F. Since $g(x), h(x) \in F[x]$, therefore we can write

$$g(x)=\frac{g_0(x)}{a},\ h(x)=\frac{h_0(x)}{b},$$

where $a, b \in R$ and where $g_0(x), h_0(x) \in R[x]$.

Also $g_0(x)=\alpha g_1(x)$, $h_0(x)=\beta h_1(x)$, where

$\alpha=c(g_0)$, $\beta=c(h_0)$ and where $g_1(x)$, $h_1(x)$ are primitive members of $R[x]$. [See theorem 3].

Thus $f(x)=\frac{\alpha\beta}{ab}g_1(x)\,h_1(x)$

$\Rightarrow ab\,f(x)=\alpha\beta\,g_1(x)\,h_1(x)$. ...(1)

Since $g_1(x)$ and $h_1(x)$ are both primitive members of $R[x]$, therefore $g_1(x)\,h_1(x)$ is also a primitive member of $R[x]$. Therefore from (1), we conclude that $f(x)$ and $g_1(x)\,h_1(x)$ are associates in $R[x]$. (See theorem 3). Thus

$f(x)=u g_1(x)\,h_1(x)$ where u is a unit in $R[x]$ and so a unit of R.

Now $u \in R$, $g_1(x) \in R[x] \Rightarrow ug_1(x)$ is a member of $R[x]$. Also $h_1(x)$ is a member of $R[x]$. Hence $f(x)$ can be factored as the product of two polynomials with coefficients in R.

Note. Let I be the ring of integers. Then I is a Euclidean ring and so a unique factorization domain. The field of quotients of I is the field of rational numbers. If in the above theorem we take I in place of R, then the statement of the theorem is as follows :

If the primitive polynomial $f(x) \in \mathrm{I}[x]$ can be factored as the product of two polynomials having rational coefficients it can be factored as the product of two polynomials having integer coefficients.

Fos its proof simply replace R by I or say, 'let R be the ring of integers.'

Theorem 8. *Let F be the field of quotients of a unique factorization domain R. If $f(x) \in R[x]$ is both primitive and irreducible as an element of $R[x]$, then it is irreducible as an element of $F[x]$. Conversely, if the primitive element $f(x)$ in $R[x]$ is irreducible as an element of $F[x]$, it is also irreducible as an element of $R[x]$.*

Proof. Let $f(x)$ be a primitive member of $R[x]$. Suppose $f(x)$ is irreducible in $R[x]$ but is reducible in $F[x]$. Since F is a field and $f(x)$ is reducible in $F[x]$, therefore we must have $f(x)=g(x)\,h(x)$, where $g(x)$, $h(x)$ are in $F[x]$ and are of positive degree. Now we can write

$$g(x)=\frac{g_0(x)}{a},\ h(x)=\frac{h_0(x)}{b},$$

where $a, b \in R$, and where $g_0(x), h_0(x) \in R[x]$.

Also $g_0(x)=\alpha\, g_1(x)$, $h_0(x)=\beta\, h_1(x)$,

where $\alpha=c(g_0)$, $\beta=c(h_0)$, and where $g_1(x)$, $h_1(x)$ are primitive members of $R[x]$. [See theorem 3]. Thus

$$f(x)=\frac{\alpha\beta}{ab}\, g_1(x)\, h_1(x)$$

$$\Rightarrow ab\, f(x)=\alpha\beta\, g_1(x)\, h_1(x). \qquad \ldots(1)$$

Since $g_1(x)$ and $h_1(x)$ are both primitive members of $R[x]$, therefore $g_1(x)\,h_1(x)$ is also a primitive member of $R[x]$. Therefore from (1), we conclude that $f(x)$ and $g_1(x)\,h_1(x)$ are associates in $R[x]$. (See theorem 3). Thus $f(x)=ug_1(x)\,h_1(x)$ where u is a unit in $R[x]$ and so a unit in R.

Let $ug_1(x)=g_2(x)$. Then

$$f(x)=g_2(x)\,h_1(x), \text{ where } g_2(x), h_1(x) \in R[x].$$

We have

deg $g_1(x)=$ deg $g(x)$, and deg $h_1(x)=$ deg $h(x)$.

Thus deg $g_1(x) > 0$, deg $h_1(x) > 0$.

Therefore neither $g_1(x)$ nor $h_1(x)$ is a unit in $R[x]$.

Thus $f(x)=g_1(x)\,h_1(x)$ is a proper factorization of $f(x)$ in $R[x]$. This contradicts the given statement that $f(x)$ is irreducible in $R[x]$. Hence $f(x)$ must be irreducible in $F[x]$.

Converse. Suppose $f(x)$ is a primitive member of $R[x]$ and is irreducible as an element of $F[x]$. Then to prove that $f(x)$ is also irreducible as an element of $R[x]$. Let

$$f(x)=g(x)\,h(x), \text{ where } g(x),\ h(x) \in R[x].$$

Then $f(x)$ will be irreducible in $R[x]$ if one of $g(x)$ or $h(x)$ is a unit in $R[x]$ *i.e.*, a unit in R.

Now $g(x),\ h(x) \in R[x]$ can also be treated as

$$g(x),\ h(x) \in F[x].$$

Since $f(x)$ is irreducible as an element of $F[x]$, therefore one of $g(x)$ or $h(x)$ must be of degree 0. Suppose deg $g(x)=0$. Then $g(x)$ is a constant polynomial.

Let $g(x)=k \in R$. Then

$$f(x)=kh(x).$$

Now $f(x)$ is a primitive member of $R[x]$. Therefore $c(f)$ is a unit in R. If k is not a unit in R, then content of $kh(x)$ cannot be a unit in R and so it cannot be equal to $c(f)$. Hence k must be a unit in R. Consequently $f(x)$ is irreducible as an element of $R[x]$.

This completes the proof of the theorem.

Theorem 9. *Let F be the field of quotients of a unique factorization domain R. If $f_1(x),\ f_2(x)$ are two primitive members of $R[x]$ and are associates in $F[x]$, then they are also associates in $R[x]$.*

Proof. Since $f_1(x),\ f_2(x)$ are associates in $F[x]$, therefore we have

$$f_1(x)=kf_2(x) \text{ where } 0\neq k \in F$$

i.e., k is a unit in $F[x]$.

Note that the only units of $F[x]$ are the non-zero elements of F.

Now F is the field of quotients of R. Therefore

$$0\neq k \in F \Rightarrow k=\frac{g}{h} \text{ where } g,\ 0\neq h \in R.$$

$$\therefore \quad f_1(x)=\frac{g}{h} f_2(x)$$

$$\Rightarrow hf_1(x)=g f_2(x).$$

Since $h, g \in R$ and $f_1(x), f_2(x)$ are primitive members of $R[x]$, therefore by theorem 3, $f_1(x)$ and $f_2(x)$ are associates in $R[x]$.

Theorem 10. *If R is a unique factorization domain and if $p(x)$ is a primitive polynomial in $R[x]$, then it can be factored in a unique way as the product of irreducible elements in $R[x]$.* **(Meerut 1970)**

Hence show that the polynomial ring $R[x]$ over a unique factorization domain R is itself a unique factorization domain.

(Delhi 1970; Calicut 75)

Proof. Let F be the field of quotients of a unique factorization domain R. Let $p(x)$ be a primitive member of $R[x]$. We can regard $p(x)$ as a member of $F[x]$. Since F is a field, therefore $F[x]$ is a unique factorization domain. Recall that the ring of polynomials over a field is a unique factorization domain. Therefore $p(x) \in F[x]$ can be factored as

$$p(x)=p_1(x)\, p_2(x) \ldots p_k(x), \text{ where}$$

$p_1(x), p_2(x), \ldots, p_k(x)$ are irreducible polynomials in $F[x]$. Now each $p_i(x)$, $1 \leqslant i \leqslant k$, can be written as $p_i(x)=\dfrac{f_i(x)}{a_i}$, where $a_i \in R$ and $f_i(x) \in R[x]$. Further $f_i(x)$ can be written as

$$f_i(x)=c_i\, q_i(x),$$

where $c_i \in R$ and $q_i(x)$ is a primitive member of $R[x]$. Thus each $p_i(x)=\dfrac{c_i}{a_i}\, q_i(x)$, where c_i, a_i R and $q_i(x)$ is a primitive member of $R[x]$. Since $p_i(x)$ is irreducible in $F[x]$, therefore $q_i(x)$ must also be irreducible in $F[x]$. Now $q_i(x)$ is a primitive member of $R[x]$ and $q_i(x)$ is irreducible in $F[x]$. Therefore, by the converse part of theorem 8, $q_i(x)$ is irreducible in $R[x]$.

Now $p(x)=p_1(x)\, p_2(x) \ldots p_k(x)$

$$=\frac{c_1\, c_2 \ldots c_k}{a_1\, a_2 \ldots a_k}\, q_1(x)\, q_2(x) \ldots q_k(x).$$

$$\therefore \quad a_1 a_2 \ldots a_k p(x)=c_1\, c_2 \ldots c_k\, q_1(x)\, q_2(x) \ldots q_k(x). \qquad \ldots(1)$$

Since $q_1(x), \ldots, q_k(x)$ are all primitive members of $R[x]$, therefore

$$q_1(x)\, q_2(x) \ldots q_k(x)$$

also a primitive member of $R[x]$. Further $p(x)$ is primitive.

Therefore from the relation (1), we conclude with the help of theorem 3 that $p(x)$ and $q_1(x)\, q_2(x) \ldots q_k(x)$ are associates in $R[x]$. Therefore $p(x) = u\, q_1(x) q_2(x) \ldots q_k(x)$
where u is some unit in $R[x]$ and hence in R.

If $q_1(x)$ is irreducible in $R[x]$ then $uq_1(x)$ is also irreducible in $R[x]$. If we simply replace $uq_1(x)$ by $q_1(x)$, then we get

$$p(x) = q_1(x)\, q_2(x) \ldots q_k(x).$$

Thus we have factored $p(x)$ in $R[x]$ as a product of irreducible elements.

Now to show that the above factorization of $p(x)$ is unique upto the order and associates of irreducible elements.

Let $p(x) = r_1(x)\, r_2(x) \ldots r_m(x)$, where the $r_i(x)$ are irreducible in $R[x]$. Since $p(x)$ is primitive, therefore each $r_i(x)$ must be primitive. Consequently by theorem 8, each $r_i(x)$ must be irreducible in $F[x]$. But $F[x]$ is a unique factorization domain. Therefore $p(x) \in F[x]$ can be uniquely expressed as the product of irreducible elements of $F[x]$. Hence the $r_i(x)$ and the $q_i(x)$ regarded as the elements of $F[x]$ are equal (upto associates) in some order.

Since $r_i(x)$ and $q_i(x)$ are primitive members of $R[x]$ and are associates in $F[x]$, therefore, by theorem 9, they are also associates in $R[x]$. Thus $p(x)$ has a unique factorization as a product of irreducible elements of $R[x]$.

Now we are in a position to prove that if R is a unique factorization domain, then so is $R[x]$.

Let $f(x) \in R[x]$ be arbitrary. Then we can write $f(x)$ in a unique way as $f(x) = cg(x)$ where $c \in R$ and $g(x)$ is a primitive member of $R[x]$.

Now by the above discussion $g(x)$ can be uniquely expressed as the product of irreducible elements of $R[x]$. What about c? Let $c = h_1(x)\, h_2(x) \ldots h_s(x)$, where $h_1(x), \ldots, h_s(x) \in R[x]$. We have

$$0 = \deg c = \deg h_1(x) + \deg h_2(x) + \ldots + \deg h_s(x)$$
$$\Rightarrow \text{each } h_i(x) \text{ must be of degree } 0$$
$$\Rightarrow \text{each } h_i(x) \text{ is an element of } R.$$

Thus the only factorizations of c as an element of $R[x]$ are those it had as an element of R. In particular if $\alpha \in R$ is irreducible, then $\alpha \in R[x]$ is also irreducible. But R is a unique factorization domain. Therefore $c \in R$ can be uniquely expressed as the product of irreducible elements of R and hence of $R[x]$.

Finally, we conclude that $f(x) = cg(x)$ can be uniquely expressed as the product of irreducible elements of $R[x]$. Hence $R[x]$ is a unique factorization domain.

Corollary 1. *If R is a unique factorization domain then so is $R[x_1, \ldots, x_n]$.*

Proof. If R is a unique factorization domain, then we know that $R_1 = R[x_1]$ is a unique factorization domain. Now R_1 is a unique factorization domain implies that

$$R_2 = R_1[x_2] = R[x_1, x_2]$$

is a unique factorization domain. Continuing this process a finite number of times we conclude that $R[x_1, \ldots, x_n]$ is a unique factorization domain.

Corollary 2. *If F is a field, then $F[x_1, x_2, \ldots, x_n]$ is a unique factorization domain.*

Proof. If F is a field, then we know that $F_1 = F[x_1]$ is a unique factorization domain. Now F_1 is a unique factorization domain implies that $F_2 = F_1[x_2] = F[x_1, x_2]$ is a unique factorization domain. Continuing this process a finite number of times we conclude that $F[x_1, \ldots, x_n]$ is a unique factorization domain.

Eisenstein's Criterion of Irreducibility.

Theorem 11. *Let F be the field of quotients of a unique factorization domain R. If*

$$f(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n \in R[x]$$

and p is a prime element of R such that

$$p \mid a_0, p \mid a_1, p \mid a_2, \ldots, p \mid a_{n-1}$$

whereas p is not a divisor of a_n and p^2 is not a divisor of a_0, then $f(x)$ is irreducible in $F[x]$. **(Banaras 1972)**

Proof. Without loss of generality we may take $f(x)$ to be primitive, for taking out the greatest common factor of its coefficients does not disturb the hypothesis, since p is not a divisor of a_n. Now suppose that $f(x)$ is reducible in $F[x]$. Then $f(x)$ can be factored as the product of two ploynomials of positive degree in $F[x]$. Therefore by Gauss lemma $f(x)$ can be factored as the product of two polynomials of positive degree in $R[x]$. Thus if we assume that $f(x)$ is reducible in $F[x]$, then

$$f(x) = a_0 + a_1x + \cdots + a_nx^n = (b_0 + b_1x + \ldots + b_rx^r)(c_0 + c_1x + \ldots + c_sx^s) \quad \ldots(1)$$

where the b's and c's are elements of R and where $r > 0$ and $s > 0$.

We have from (1), $a_0 = b_0c_0$.

Since p is a prime element of R, therefore

$$p \mid a_0 \Rightarrow p \mid b_0 \text{ or } p \mid c_0.$$

Since p^2 is not a divisor of a_0, therefore p cannot divide both b_0 and c_0. Suppose that

$$p \mid b_0 \text{ and } p \text{ is not a divisor of } c_0.$$

If p is a divisor of all the coefficients $b_0, b_1, \ldots, b_r$, then from (1) we see that p is a divisor of all the coefficients of $f(x)$. But p is not a divisor of a_n. Therefore not all the coefficients $b_0, b_1, \ldots, b_r$ can be divisible by p. Let b_k, where $k \leqslant r$, be the first b which is not divisible by p. Then each of $b_0, b_1, \ldots, b_{k-1}$ is divisible by p and b_k is not divisible by p.

Also $k < n$, since $r < n$.

Now from (1), we have

$$a_k = b_k c_0 + b_{k-1} c_1 + b_{k-2} c_2 + \ldots + b_0 c_k$$

$$\Rightarrow b_k c_0 = a_k - b_{k-1} c_1 - b_{k-2} c_2 - \ldots - b_0 c_k. \qquad \ldots(2)$$

Now $k < n$. Therefore $p \mid a_k$. Also $p \mid b_{k-1}, b_{k-2}, \ldots, b_0$.

Therefore from (2), we have

$$p \mid b_k c_0$$

$$\Rightarrow p \mid b_k \text{ or } p \mid c_0, \text{ since } p \text{ is a prime element of } R.$$

But this is nonsense because according to our initial assumptions p is neither a divisor of b_k nor a divisor of c_0. Hence $f(x)$ must be irreducible in $F[x]$. This completes the proof of the theorem.

Note. In the above theorem if we take the ring of integers I in place of the unique factorization domain R, then the field of quotients of I is the field of rational numbers. The statement of the theorem will be as follows :

Let $f(x) = a_0 + a_1 x + \ldots + a_n x^n$ *be a polynomial with integer coefficients. If p is a prime number such that*

$$p \mid a_0, \ p \mid a_1, \ \ldots, \ p \mid a_{n-1}$$

whereas p is not a divisor of a_n *and* p^2 *is not a divisor of* a_0, *then* $f(x)$ *is irreducible over the field of rational numbers.*

(Madras 1983; Meerut 74)

There will be no difference in proof.

Solved Examples

Ex. 1. *If p is a prime number, prove that the polynomial* $x^n - p$ *is irreducible over the field of rational numbers.*

Solution. Let $f(x) = x^n - p$

$$= p + 0x + 0x^2 + \ldots + 0x^{n-1} + 1x^n.$$

Then $f(x)$ is a polynomial with integer coefficients.

Now p is a prime number.

We see that p divides each of the coefficients of $f(x)$ except the coefficient 1 of the last term x^n. Also p^2 is not a divisor of the constant term $-p$. Hence by Eisenstein's criterion of irreducibility $f(x)$ is irreducible over the field of rational numbers.

Ex. 2. *Show that the polynomial x^2-3 is irreducible over the field of rational numbers.*

Solution. Let $f(x)=x^2-3=-3+0x+1x^2$.

Now $f(x)$ is a polynomial with integer coefficients. Also 3 is a prime number such that 3 divides each of the coefficients of $f(x)$ except the coefficient 1 of the last term x^2. Also 3^2 is not a divisor of the constant term -3. Hence by Eisenstein's criterion of irreducibility $f(x)$ is irreducible over the field of rational numbers.

Ex. 3. *Prove that the polynomial $1+x+\dots+x^{p-1}$, where p is a prime number, is irreducible over the field of rational numbers.*

(Banaras 1963; Madras 83)

Solution. Let $f(x)=1+x+\dots+x^{p-1}$.

Multiplying both sides by $x-1$, we get

$$(x-1)\,f(x)=(x-1)\,(x^{p-1}+x^{p-2}+\dots+x+1)$$

$$\Rightarrow (x-1)\,f(x)=x^p-1.$$

Putting $x-1=y$ or $x=y+1$ on both sides, we get

$$yf\,(y+1)=(y+1)^p-1$$

$$={y^p}+{}^pC_1y^{p-1}+{}^pC_2y^{p-2}+\dots+{}^pC_{p-1}\,y+1-1,$$

expanding by binomial theorem

$$=y^p+{}^pC_1\,y^{p-1}+{}^pC_2y^{p-2}+\dots+{}^pC_{p-1}\,y$$

$$=y\,[y^{p-1}+{}^pC_1y^{p-2}+{}^pC_2y^{p-3}+\dots+{}^pC_{p-2}\,y+{}^pC_{p-1}]$$

$$\therefore\quad f(y+1)=y^{p-1}+{}^pC_1y^{p-2}+{}^pC_2y^{p-3}+\dots+{}^pC_{p-2}y+{}^pC_{p-1}.$$

$$\text{Now}\quad {}^pC_r=\frac{p\,(p-1)\,(p-2)\dots(p-r+1)}{r\,!},\ 1\leqslant r\leqslant p-1.$$

Obviously ${}^pC_r$ is divisible by p for each $1 \leqslant r \leqslant p-1$. Note that p is given to be a prime integer.

Now $f(y+1)$ is a polynomial with integer coefficients. Also p is a prime number such that p divides each of the coefficients of $f(y+1)$ except the coefficient of y^{p-1} which is 1. Also p^2 is not a divisor of the constant term which is equal to ${}^pC_{p-1}=p$. Therefore by Eisenstein's criterion of irreducibility $f(y+1)$ is irreducible over the field of rational numbers. Therefore $f(x)$ is irreducible over the field of rational numbers. Note that $y=x-1$.

Ex. 4. *Show that the polynomial $x^4+x^3+x^2+x+1$ is irreducible over the field of rational numbers.*

Solution. The number 5 is a prime number. So proceed as in Ex. 3 by taking $p=5$.

Ex. 5. *Let R be a unique factorization domain. Then show that every prime element in R generates a prime ideal.*

Solution. Let p be a prime element of a unique factorization domain R. Let $S=(p)$ be the ideal of R generated by p. Then to show that S is a prime ideal.

Suppose ab is an element of S where $a, b \in R$. We have

$$
\begin{aligned}
ab \in S &\Rightarrow ab=kp \text{ for some } k \in R\\
&\Rightarrow p \mid ab\\
&\Rightarrow p \mid a \text{ or } p \mid b \qquad [\because \ p \text{ is a prime element of } R]\\
&\Rightarrow a=sp \text{ or } b=tp \text{ for some } s, t \in R\\
&\Rightarrow a \in (p) \text{ or } b \in (p)\\
&\Rightarrow (p) \text{ is a prime ideal of } R.
\end{aligned}
$$

Exercises

1. Examine whether the polynomial x^3+3x^2+x-4 is irreducible over (i) the field of integers modulo 5, (ii) the field of integers modulo 7. **(Bombay 1970)**

2. Let $p(x)$ be an irreducible polynomial over a field F. Prove that the ideal generated by $p(x)$ in $F[x]$ is a maximal ideal. **(Kurukshetra 1971)**

3. Show that if F is a unique factorization domain then $F[x]$ is also a unique factorization domain. **(Delhi 1970)**

4. Prove that a Euclidean ring is necessarily a principal ideal ring with unity. Give two examples of such a ring. **(Banaras 1968)**

5. Prove that I $[\sqrt{2}]$, the set of real numbers $a+b\sqrt{2}$ where a, b are integers is a Euclidean ring. **(Kurukshetra 1970)**

6. Prove that the ring of polynomials over a field F is a Euclidean domain. Hence or otherwise deduce that any non-zero polynomials $f(x), g(x) \in F[x]$ have g.c.d. **(Meerut 1976)**

6

Vector Spaces

So far we have studied groups and rings. Now we shall study another important algebraic structure known as **vector space**. Before giving the definition of a vector space we shall make a distinction between **internal and external compositions.**

Let A be any set. If $a * b \in A \;\forall a, b \in A$, and $a * b$ is unique then $*$ is said to be an **internal composition** in the **set** A. Here a and b are both elements of the set A.

Let V and F be any two sets. If $a \circ \alpha \in V$ for all $a \in F$ and for all $\alpha \in V$ and $a \circ \alpha$ is unique, then $\circ$ is said to be an **external composition** in V over F. Here a is an element of the set F and α is an element of the set V and the resulting element $a \circ \alpha$ is an element of the set V.

§ 1. Vector space. Definition. **(Madras 1977; Banaras 70; Rohilkhand 80; Kolhapur 73; Poona 73; Meerut 76)**

Let $(F, +, \cdot)$ *be a field. The elements of F will be called* **scalars**. *Let V be a non-empty set whose elements will be called* **vectors**. *Then V is a vector space over the field F, if*

1. *There is defined an internal composition in V called* **addition of vectors** *and denoted by* '$+$'. *Also for this composition V is an abelian group.*

2. *There is an external composition in V over F called* **scalar multiplication** *and denoted multiplicatively i.e.,* $a\alpha \in V$ *for all* $a \in F$ *and for all* $\alpha \in V$. *In other words V is closed with respect to scalar multiplication.*

3. *The two compositions i.e., scalar multiplication and addition of vectors satisfy the following postulates :*

(*i*) $a(\alpha+\beta)=a\alpha+a\beta \;\forall a\in F$ *and* $\forall \alpha, \beta \in V$.

(*ii*) $(a+b)\,\alpha=a\alpha+b\alpha \;\forall a, b\in F$ *and* $\forall \alpha \in V$.

(*iii*) $(ab)\,\alpha=a(b\alpha) \;\forall a, b\in F$ *and* $\forall \alpha \in V$.

(*iv*) $1\alpha=\alpha \;\forall \alpha \in V$ *and* 1 *is the unity element of the field F.*

When V is a vector space over the field F, we shall say that $V(F)$ is a vector space. If the field F is understood we can simply say that V is a vector space.

In the above definition of a vector space V over the field F, we have denoted the addition of vectors by the symbol '+'. This symbol also denotes the addition composition of the field F, *i.e.*, addition of scalars. There should be no confusion about the two compositions though we have used the same symbol to denote each of them. If $\alpha, \beta \in V$, then $\alpha+\beta$ represents addition in V *i.e.*, addition of vectors. If $a, b \in F$ then $a+b$ represents addition of scalars *i e.*, addition in the field F. Similarly there should be no confusion in multiplication of scalars *i.e.*, multiplication of the elements of F and in scalar multiplication *i.e.*, multiplication of an element of V by an element of F. If $a, b \in F$, then ab represents multiplication of F and $ab \in F$. If $a \in F$, and $\alpha \in V$, then $a\alpha$ represents scalar multiplication and $a\alpha \in V$. Since $1 \in F$ and $\alpha \in V$, therefore 1α represents scalar multiplication. Again $a\alpha \in V$, $a\beta \in V$, therefore $a\alpha+a\beta$ represents addition of vectors and thus $a\alpha+a\beta$ is an element of V. Further $a \in F$ and $\alpha+\beta \in V$, therefore $a(\alpha+\beta)$ represents scalar multiplication and we have $a(\alpha+\beta) \in V$.

Note 1. For V to be an abelian group with respect to addition . vectors, we must have the following conditions satisfied :

(i) $\alpha+\beta \in V$ for all $\alpha, \beta \in V$.

(ii) $\alpha+\beta=\beta+\alpha$ for all $\alpha, \beta \in V$.

(iii) $\alpha+(\beta+\gamma)=(\alpha+\beta)+\gamma$ for all $\alpha, \beta, \gamma \in V$.

(iv) There exists an element $\mathbf{0} \in V$ such that

$$\mathbf{0}+\alpha=\alpha \ \forall \ \alpha \in V.$$

This element $\mathbf{0} \in V$ will be called the **zero vector**. It is the additive identity in V.

(v) To every vector $\alpha \in V$ there exists a vector $-\alpha \in V$ such that $-\alpha+\alpha=\mathbf{0}$. Thus each vector should possess additive inverse. The vector $-\alpha$ is called the negative of the vector α.

Note 2. Since $(V, +)$ is an abelian group therefore all the properties of an abelian group will hold in V. A few of them are as follows:

(i) $\alpha+\beta=\alpha+\gamma \Rightarrow \beta=\gamma$ (left cancellation law)

(ii) $\beta+\alpha=\gamma+\alpha \Rightarrow \beta=\gamma$ (right cancellation law)

(iii) $\alpha+\beta=\mathbf{0} \Rightarrow \alpha=-\beta$ and $\beta=-\alpha$.

(iv) $-(\alpha+\beta)=-\alpha-\beta$ where by $\alpha-\beta$ we mean $\alpha+(-\beta)$.

(v) The additive identity **0** will be unique.

(vi) The additive inverse of each vector will be unique.

(vii) If $\alpha+\beta=\gamma$, then $\alpha+\beta-\gamma=\mathbf{0}$.

Note 3. There should also be no confusion about the use of the word vector. Here by vector we do not mean the vector quantity which we have defined in vector algebra as a directed line segment. Here we shall call the elements of the set V as vectors.

Note 4. In a vector space we shall be dealing with two types of zero elements. One is the zero vector and the other is the zero element of the field F *i.e.*, the 0 scalar. To distinguish between the two, we shall use the zero letter in bold type to represent the zero vector. However the students may use the same symbol 0 to denote the zero vector as well as the zero scalar. There will be no confusion in this use. The use of 0 will itself tell whether it stands for zero vector or for zero scalar.

Note 5. We shall use the lower case Greek letters α, β, γ etc. to denote vectors *i.e.*, the elements of V and the lower case Latin letters a, b, c etc. to denote the scalars *i.e.*, the elements of the field F.

Example 1. *A field K can be regarded as a vector space over any subfield F of K.* **(Meerut 1969; Banaras 69; Kumayon 78)**

Here K is the set of vectors. Addition of vectors is the addition composition in the field K. Since K is a field, therefore $(K, +)$ is an abelian group. Further the elements of the subfield F constitute the set of scalars. The composition of scalar multiplication is the multiplication composition in the field K. K is a field, therefore $a\alpha \in K \; \forall \; a \in F$ and $\forall \; \alpha \in K$ because both a and α are elements of K. If 1 is the unity element of K, then 1 is also the unity element of the subfield F. We make the following observations.

(i) $a(\alpha+\beta)=a\alpha+a\beta \; \forall \; a \in F$ and $\forall \; \alpha, \beta \in K$. This result follows from the left distributive law in K.

(ii) $(a+b)\,\alpha=a\alpha+b\alpha \; \forall \; a, b \in F$ and $\forall \; \alpha \in K$. This result is a consequence of the right distributive law in K.

(iii) $(ab)\,\alpha=a(b\alpha) \; \forall \; a, b \in F$ and $\forall \; \alpha \in K$. This result is a consequence of associativity of multiplication in K.

(iv) $1\alpha=\alpha \; \forall \; \alpha \in K$ and 1 is the unity element of the subfield F.

Since 1 is also the unity element of the field K, therefore $1\alpha=\alpha \ \forall \ \alpha \in K$. Hence $K(F)$ is a space.

Note 1. If F is any field, then F itself is a vector space over the field F.

Note 2. If **C** is the field of complex numbers and **R** is the field of real numbers, then **C** is a vector space over **R** because **R** is a subfield of **C**. But **R** is not a vector space over **C**. Here **R** is not closed with respect to scalar multiplication. For example, $2 \in \mathbf{R}$ and $3+4i \in \mathbf{C}$ and $(3+4i)\ 2 \notin \mathbf{R}$.

Example 2. *The set V of all $m \times n$ matrices with their elements as real numbers is a vector space over the field F of real numbers with respect to addition of matrices as addition of vectors and multiplication of a matrix by a scalar as scalar multiplication.*

As in groups, we can prove that V is an abelian group with respect to addition of matrices. The null matrix O of the type $m \times n$ is the additive identity of this abelian group.

If $a \in F$ and $\alpha \in V$ (*i.e.*, α is a matrix of the type $m \times n$ with elements as real numbers), then $a\alpha \in V$ because $a\alpha$ is also a matrix of the type $m \times n$ with elements as real numbers. Therefore V is closed with respect to scalar multiplication Also from our study of matrices we observe that

(i) $a(\alpha+\beta)=a\alpha+a\beta \ \forall \ a \in F$ and $\forall \ \alpha, \beta \in V$.

(ii) $(a+b)\ \alpha=a\alpha+b\alpha \ \forall \ a, b \in F$ and $\forall \ \alpha \in V$.

(iii) $(ab)\ \alpha=a(b\alpha) \quad \forall \ a, b \in F$ and $\forall \ \alpha \in V$.

(iv) $1\alpha=\alpha \ \forall \ \alpha \in V$ where 1 is the unity element of the field F of real numbers.

Hence $V(F)$ is a vector space.

Note. If V is the set of all $m \times n$ matrices with their elements as rational numbers and F is the field of real numbers, then V will not be closed with respect to scalar multiplication. For $\sqrt{7} \in F$ and if $\alpha \in V$, then $\sqrt{7}\alpha \notin V$ because the elements of the matrix $\sqrt{7}\alpha$ will not be rational numbers. Therefore $V(F)$ will not be a vector space.

Example 3. *The vector space of all ordered n-tuples over a field F.*

Let F be a field. An ordered set $\alpha=(a_1, a_2, a_3, \ldots, a_n)$ of n elements of F is called an *n-tuple* over F. Let V be the totality of all ordered *n-tuples over F i.e.*, let

$$V=\{(a_1, a_2, \ldots, a_n) : a_1, a_2, a_3, \ldots, a_n \in F\}.$$

Now we shall give a vector space structure to V over the field F. For this we define equality of two n-tuples, addition of two n-tuples and multiplication of an n-tuple by a scalar as follows :

Equality of two n-tuples. Two elements $\alpha=(a_1, a_2, \ldots, a_n)$ and $\beta=(b_1, b_2, \ldots, b_n)$ of V are said to be equal if and only if

$$a_i=b_i \text{ for each } i=1, 2, \ldots, n.$$

Addition composition in V. We define

$$\alpha+\beta=(a_1+b_1, a_2+b_2, \ldots, a_n+b_n)$$
$$\forall\ \alpha=(a_1, a_2, \ldots, a_n),\ \beta=(b_1, b_2, \ldots, b_n)\in V.$$

Since $a_1+b_1, a_2+b_2, \ldots, a_n+b_n$ are all elements of F, therefore $\alpha+\beta \in V$ and thus V is closed with respect to addition of n-tuples.

Scalar multiplication composition in V over F. We define

$a\alpha=(aa_1, aa_2, \ldots, aa_n)\ \forall\ a\in F, \alpha=(a_1, a_2, \ldots, a_n)\in V.$

Since $aa_1, aa_2, \ldots, aa_n$ are all elements of F, therefore $a\alpha\in V$ and thus V is closed with respect to scalar multiplication.

Now we shall see that V is a vector space for these two compositions.

Associativity of addition in V. We have

$$\begin{aligned}
&(a_1, a_2, \ldots, a_n)+[(b_1, b_2, \ldots, b_n)+(c_1, c_2, \ldots, c_n)]\\
&=(a_1, a_2, \ldots, a_n)+(b_1+c_1, b_2+c_2, \ldots, b_n+c_n)\\
&=(a_1+[b_1+c_1], a_2+[b_2+c_2], \ldots, a_n+[b_n+c_n])\\
&=([a_1+b_1]+c_1, [a_2+b_2]+c_2, \ldots, [a_n+b_n]+c_n)\\
&=(a_1+b_1, a_2+b_2, \ldots, a_n+b_n)+(c_1, c_2, \ldots, c_n)\\
&=[(a_1, a_2, \ldots, a_n)+(b_1, b_2, \ldots, b_n)]+(c_1, c_2, \ldots, c_n).
\end{aligned}$$

Commutativity of addition in V. We have

$$\begin{aligned}
&(a_1, a_2, \ldots, a_n)+(b_1, b_2, \ldots, b_n)=(a_1+b_1, a_2+b_2, \ldots, a_n+b_n)\\
&=(b_1+a_1, b_2+a_2, \ldots, b_n+a_n)=(b_1, b_2, \ldots, b_n)+(a_1, a_2, \ldots, a_n).
\end{aligned}$$

Existence of additive identity in V. We have

$(0, 0, \ldots, 0)\in V$. Also if $(a_1, a_2, \ldots, a_n)\in V$, then

$$\begin{aligned}
(a_1, a_2, \ldots, a_n)+(0, 0, \ldots, 0)&=(a_1+0, a_2+0, \ldots, a_n+0)\\
&=(a_1, a_2, \ldots, a_n).
\end{aligned}$$

$\therefore$ $(0, 0, \ldots, 0)$ is the additive identity in V.

Existence of additive inverse of each element of V. If $(a_1, a_2, \ldots, a_n)\in V$, then $(-a_1, -a_2, \ldots, -a_n)\in V$.

Also we have $(-a_1, -a_2, \ldots, -a_n)+(a_1, a_2, \ldots, a_n)$

$=(-a_1+a_1, -a_2+a_2, \ldots, -a_n+a_n)=(0, 0, 0, \ldots, 0).$

$\therefore$ $(-a_1, -a_2, \ldots, -a_n)$ is the additive inverse of $(a_1, a_2, \ldots, a_n)$.

Thus V is an abelian group with respect to addition. Further we observe that

1. If $a \in F$ and $\alpha=(a_1, a_2, \ldots, a_n)$, $\beta=(b_1, b_2, \ldots, b_n) \in V$, then
$$a(\alpha+\beta)=a(a_1+b_1, a_2+b_2, \ldots, a_n+b_n)$$
$$=(a[a_1+b_1], a[a_2+b_2], \ldots, a[a_n+b_n])$$
$$=(aa_1+ab_1, aa_2+ab_2, \ldots, aa_n+ab_n)$$
$$=(aa_1, aa_2, \ldots, aa_n)+(ab_1, ab_2, \ldots, ab_n)$$
$$=a(a_1, a_2, \ldots, a_n)+a(b_1, b_2, \ldots, b_n)=a\alpha+a\beta.$$

2. If $a, b \in F$ and $\alpha=(a_1, a_2, \ldots, a_n) \in V$, then
$$(a+b)\alpha=([a+b]a_1, [a+b]a_2, \ldots, [a+b]a_n)$$
$$=(aa_1+ba_1, aa_2+ba_2, \ldots aa_n+ba_n)$$
$$=(aa_1, aa_2, \ldots, aa_n)+(ba_1, ba_2, \ldots, ba_n)$$
$$=a(a_1, a_2, \ldots, a_n)+b(a_1, a_2, \ldots, a_n)=a\alpha+b\alpha.$$

3. If $a, b \in F$ and $\alpha=(a_1, a_2, \ldots, a_n) \in V$, then
$$(ab)\alpha=([ab]a_1, [ab]a_2, \ldots, [ab]a_n)=(a[ba_1], a[ba_2], \ldots, a[ba_n])$$
$$=a(ba_1, ba_2, \ldots, ba_n)=a[b(a_1, a_2, \ldots, a_n)]=a(b\alpha).$$

4. If 1 is the unity element of F and $\alpha=(a_1, a_2, \ldots, a_n) \in V$, then $1\alpha=(1a_1, 1a_2, \ldots, 1a_n)=(a_1, a_2, \ldots, a_n)=\alpha$.

Hence V is a vector space over F. *The vector space of all ordered n-tuples over F will be denoted by $V_n(F)$. Sometimes we also denote it by F^n.* Here the zero vector *i.e.*, **0** is the n-tuple $(0, 0, \ldots, 0)$.

Note. $V_2(F)=\{(a_1, a_2) : a_1, a_2 \in F\}$ is the vector space of all **ordered pairs** over F. Similarly $V_3(F)=\{(a_1, a_2, a_3) : a_1, a_2, a_3 \in F\}$ is the vector space of all **ordered triads** over F.

Example 4. *The vector space of all polynomials over a field F.*

Let $F[x]$ denote the set of all polynomials in indeterminate x over a field F. Then $F[x]$ is a vector space over F with respect to addition of two polynomials as addition of vectors and the product of a polynomial by a constant polynomial (*i.e.*, by an element of F) as scalar multiplication.

Example 5. *The vector shace of all real valued continuous (differentiable or integrable) functions defined in some interval* [0, 1].

Let V denote the set of all real valued continuous functions of x defined in the interval [0, 1]. Then V is a vector space over the field R of real numbers with vector addition and scalar multiplication defined as below :
$$(f+g)(x)=f(x)+g(x) \ \forall \ f, g \in V$$
and
$$(af)(x)=af(x) \ \forall \ a \in R, f \in V.$$

As in rings, we should first prove that V is an abelian group with respect to addition composition.

V is closed with respect to scalar multiplication since af is also a real valued continuous function in $[0, 1]$. Further we observe that

1. If $a \in R$ and $f, g \in V$, then

$$[a(f+g)](x)=a[(f+g)(x)]=a[f(x)+g(x)]=af(x)+ag(x)$$
$$=(af)(x)+(ag)(x)=(af+ag)(x).$$

$\therefore \quad a(f+g)=af+ag.$

2. If $a, b \in R$ and $f \in V$, then

$$[(a+b)f](x)=(a+b)f(x)=af(x)+bf(x)=(af)(x)+(bf)(x)$$
$$=(af+bf)(x).$$

$\therefore \quad (a+b)f=af+bf.$

3. If $a, b \in R$ and $f \in V$, then

$$[(ab)f](x)=(ab)f(x)=a[bf(x)]=a[(bf)(x)]=[a(bf)](x).$$

$\therefore \quad (ab)f=a(bf).$

4. If 1 is the unity element of R and $f \in V$, then

$$(1f)(x)=1f(x)=f(x).$$

$\therefore \quad 1f=f.$

Hence V is a vector space over R.

Exercises

1. Show that a field F may be considered as a vector space over F if scalar multiplication is identified with field multiplication.

2. Show that the complex field **C** is a vector space over the real field **R**. **(Madurai 1975 ; Kumayun 77)**

3. Let V be the set of all pairs (x, y) of real numbers, and let F be the field of real numbers. Define

$$(x, y)+(x_1, y_1)=(x+x_1, y+y_1)$$
$$c(x, y)=(cx, y)$$

Show that with these operations V is not a vector space over the field of real numbers.

4. Let $V=\mathbf{R}^2=\{(a_1, a_2) : a_1, a_2 \in \mathbf{R}\}$ and $F=\mathbf{R}$.

Define the addition and scalar multiplication in $\mathbf{R}^2$ as follows :

$$(a_1, a_2)+(b_1, b_2)=(a_1+b_1, a_2+b_2)$$

and
$$a(a_1, a_2)=(aa_1, aa_2).$$

Show that $\mathbf{R}^2$ is a vector space over $\mathbf{R}$. **(Garhwal 1976)**

5. Let V be the set of all pairs of real numbers and let F be the field of real numbers. Define

$$(x, y)+(x_1, y_1)=(3y+3y_1, -x-x_1)$$
$$c(x, y)=(3cy, -cx).$$

Verify that V, with these operations, is not a vector space over the field of real numbers.

6. Let **R** be the field of real numbers and let P_n be the set of all polynomials (of degree at most n) over the field **R**. Prove that P_n is a vector space over the field **R**.

7. How many elements are there in the vector space of polynomials of degree at most n in which the coefficients are the elements of the field $\mathbf{I}(p)$ over the field $\mathbf{I}(p)$, p being a prime number ? (Meerut 1974, 75) Ans. p^{n+1}.

§ 2. General properties of vector spaces.

Theorem 1. *Let $V(F)$ be a vector space and* $\mathbf{0}$ *be the zero vector of V. Then*

(*i*) $a\mathbf{0}=\mathbf{0}\ \forall\ a\in F$.
(I.A.S. 1974 ; Delhi 70 ; Kolhapur 79 ; Meerut 69)

(*ii*) $0\alpha=\mathbf{0}\ \forall\ \alpha\in V$. (Delhi 1970 ; I.A.S. 74 ; Meerut 69)

(*iii*) $a(-\alpha)=-(a\alpha)\ \forall\ a\in F,\ \forall\ \alpha\in V$. (Delhi 1970)

(*iv*) $(-a)\alpha=-(a\alpha)\ \forall\ a\in F,\ \forall\ \alpha\in V$. (I.A.S. 1974)

(*v*) $a(\alpha-\beta)=a\alpha-a\beta\ \forall\ a\in F$ *and* $\forall\ \alpha,\beta\in V$.

(*vi*) $a\alpha=\mathbf{0}\Rightarrow a=0$ *or* $\alpha=\mathbf{0}$. (Meerut 1968 ; I A.S. 74)

Proof. (i) We have $a\mathbf{0}=a(\mathbf{0}+\mathbf{0})$ $[\because\ \mathbf{0}=\mathbf{0}+\mathbf{0}]$

$=a\mathbf{0}+a\mathbf{0}$.

$\therefore\ \mathbf{0}+a\mathbf{0}=a\mathbf{0}+a\mathbf{0}$ $[\because\ a\mathbf{0}\in V$ and $\mathbf{0}+a\mathbf{0}=a\mathbf{0}]$

Now V is an abelian group with respect to addition.

Therefore by right cancellation law in V, we get $\mathbf{0}=a\mathbf{0}$.

(ii) We have $0\alpha=(0+0)\alpha$ $[\because\ 0=0+0]$

$=0\alpha+0\alpha$.

$\therefore\ \mathbf{0}+0\alpha=0\alpha+0\alpha$. $[\because\ 0\alpha\in V$ and $\mathbf{0}+0\alpha=0\alpha]$

Now V is an abelian group with respect to addition of vectors. Therefore by right cancellation law in V, we get $\mathbf{0}=0\alpha$.

(iii) We have $a[\alpha+(-\alpha)]=a\alpha+a(-\alpha)$

$\Rightarrow a\mathbf{0}=a\alpha+a(-\alpha)$

$\Rightarrow \mathbf{0}=a\alpha+a(-\alpha)$ $[\because\ a\mathbf{0}=\mathbf{0}]$

$\Rightarrow a(-\alpha)$ is the additive inverse of $a\alpha$

$\Rightarrow a(-\alpha)=-(a\alpha)$.

(iv) We have $[a+(-a)]\alpha=a\alpha+(-a)\alpha$

$\Rightarrow 0\alpha=a\alpha+(-a)\alpha$

$\Rightarrow \mathbf{0}=a\alpha+(-a)\alpha$

$\Rightarrow (-a)\alpha$ is the additive inverse of $a\alpha$

$\Rightarrow (-a)\alpha=-(a\alpha)$.

(v) We have $a(\alpha-\beta)=a[\alpha+(-\beta)]=a\alpha+a(-\beta)$

$=a\alpha+[-(a\beta)]$ $\qquad$ [$\because$ $a(-\beta)=-(a\beta)$]
$=a\alpha-a\beta$.

(vi) Let $a\alpha=0$ and $a\neq 0$. Then a^{-1} exists because a is a non-zero element of the field F.

$\therefore$ $a\alpha=0 \Rightarrow a^{-1}(a\alpha)=a^{-1}\,0 \Rightarrow (a^{-1}a)\,\alpha=0 \Rightarrow 1\alpha=0 \Rightarrow \alpha=0$.

Again let $a\alpha=0$ and $\alpha\neq 0$. Then to prove that $a=0$. Suppose $a\neq 0$. Then a^{-1} exists.

$\therefore$ $a\alpha=0 \Rightarrow a^{-1}(a\alpha)=a^{-1}\,0 \Rightarrow (a^{-1}a)\,\alpha=0 \Rightarrow 1\alpha=0 \Rightarrow \alpha=0$.

Thus we get a contradiction that α must be a zero vector. Therefore a must be equal to 0. Hence $\alpha\neq 0$ and $a\alpha=0 \Rightarrow a=0$.

Theorem 2. *Let $V(F)$ be a vector space. Then*

(*i*) *If $a, b \in F$ and α is a non-zero element of V, we have*

$$a\alpha=b\alpha \Rightarrow a=b.$$

(*ii*) *If $\alpha, \beta \in V$ and a is a non-zero element of F, we have*

$$a\alpha=a\beta \Rightarrow \alpha=\beta.$$ **(Meerut 1974)**

Proof. (i) We have $a\alpha=b\alpha \Rightarrow a\alpha-b\alpha=0 \Rightarrow (a-b)\,\alpha=0$.

But $\alpha\neq 0$ Therefore $(a-b)\,\alpha=0 \Rightarrow a-b=0 \Rightarrow a=b$.

(ii) We have $a\alpha=a\beta \Rightarrow a\alpha-a\beta=0 \Rightarrow a\,(\alpha-\beta)=0$.

But $a\neq 0$. Therefore $a\,(\alpha-\beta)=0 \Rightarrow \alpha-\beta=0 \Rightarrow \alpha=\beta$.

§ 3. Vector Subspaces. Definition

Let V be a vector space over the field F and $W \subseteq V$. Then W is called a subspace of V if W itself is a vector space over F with respect to the operations of vector addition and scalar multiplication in V. **(Delhi 1970 ; Madurai 78 ; Banaras 70 ; Meerut 69)**

Theorem 1. *The necessary and sufficient condition for a non-empty subset W of a vector space $V(F)$ to be a subspace of V is that W is closed under vector addition and scalar multiplication in V.*

(Kumayun 1978)

Proof. If W itself is a vector space over F with respect to vector addition and scalar multiplication in V, then W must be closed with respect to these two compositions. Hence the condition is necessary.

The condition is sufficient. Now suppose that W is a non-empty subset of V and W is closed under vector addition and scalar multiplication in V.

Let $\alpha \in W$. If 1 is the unity element of F, then $-1 \in F$. Now W is closed under scalar multiplication. Therefore

$$-1 \in F, \alpha \in W \Rightarrow (-1)\,\alpha \in W \Rightarrow -(1\alpha) \in W$$

$\Rightarrow -\alpha \in W$ $\qquad$ [$\because$ $\alpha \in W \Rightarrow \alpha \in V$ and $1\alpha=\alpha$ in V].

Thus the additive inverse of each element of W is also in W.

Now W is closed under vector addition.

Therefore $\alpha \in W, -\alpha \in W \Rightarrow \alpha+(-\alpha) \in W$

$\Rightarrow 0 \in W$ where 0 is the zero vector of V.

Hence the zero vector of V is also the zero vector of W. Since the elements of W are also the elements of V, therefore vector addition will be commutative as well as associative in W. Hence W is an abelian group with respect to vector addition. Also it is given that W is closed under scalar multiplication. The remaining postulates of a vector space will hold in W since they hold in V of which W is a subset.

Hence W itself is a vector space for the two compositions.

$\therefore$ W is a subspace of V.

Theorem 2. *The necessary and sufficient conditions for a non-empty subset W of a vector space V (F) to be a subspace of V are*

(*i*) $\alpha \in W, \beta \in W \Rightarrow \alpha-\beta \in W$.

(*ii*) $a \in F, \alpha \in W \Rightarrow a\alpha \in W$.

Proof. The conditions are necessary. If W is a subspace of V, then W is an abelian group with respect to vector addition. Therefore $\alpha \in W, \beta \in W \Rightarrow \alpha-\beta \in W$. Also W must be closed under scalar multiplication. Therefore the condition (ii) is also necessary.

The conditions are sufficient. Now suppose W is a non-empty subset of V satisfying the two given conditions. From condition (i), we have

$$\alpha \in W, \alpha \in W \Rightarrow \alpha-\alpha \in W \Rightarrow 0 \in W.$$

Thus the zero vector of V belongs to W and it will also be the zero vector of W.

Now $0 \in W, \alpha \in W \Rightarrow 0-\alpha \in W \Rightarrow -\alpha \in W$.

Thus the additive inverse of each element of W is also in W.

Again $\alpha \in W, \beta \in W \Rightarrow \alpha \in W, -\beta \in W$

$\Rightarrow \alpha-(-\beta) \in W \Rightarrow \alpha+\beta \in W$.

Thus W is closed with respect to vector addition.

Since the elements of W are also the elements of V, therefore vector addition will be commutative as well as associative in W. Hence W is an abelian group under vector addition. Also from condition (ii), W is closed under scalar multiplication. The remaining postulates of a vector space will hold in W since they hold in V which is a superset of W. Hence W is a subspace of V.

Theorem 3. *The necessary and sufficient condition for a non-empty subset W of a vector space V (F) to be a subspace of V is*

$a, b \in F$ *and* $\alpha, \beta \in W \Rightarrow a\alpha + b\beta \in W$.

(Meerut 1981; Garhwal 76)

Proof. The condition is necessary. If W is a subspace of V, then W must be closed under scalar multiplication and vector addition.

Therefore $a \in F, \alpha \in W \Rightarrow a\alpha \in W$

and $b \in F, \beta \in W \Rightarrow b\beta \in W$.

Now $a\alpha \in W, b\beta \in W \Rightarrow a\alpha + b\beta \in W$. Hence the condition is necessary.

The condition is sufficient. Now suppose W is a non-empty subset of V satisfying the given condition *i.e.*, $a, b \in F$ and

$\alpha, \beta \in W \Rightarrow a\alpha + b\beta \in W$.

Taking $a = 1, b = 1$, we see that if $\alpha, \beta \in W$ then

$1\alpha + 1\beta \in W$

$\Rightarrow \alpha + \beta \in W$. $[\because \alpha \in W \Rightarrow \alpha \in V$ and $1\alpha = \alpha$ in $V]$

Thus W is closed under vector addition.

Now taking $a = -1, b = 0$, we see that if $\alpha \in W$, then

$(-1)\,\alpha + 0\alpha \in W$ [in place of β we have taken α]

$\Rightarrow -(1\alpha) + 0 \in W \Rightarrow -\alpha \in W$.

Thus the additive inverse of each element of W is also in W.

Taking $a = 0, b = 0$, we see that if $\alpha \in W$ then

$0\alpha + 0\alpha \in W \Rightarrow 0 + 0 \in W \Rightarrow 0 \in W$.

Thus the zero vector of V belongs to W. It will also be the zero vector of W.

Since the elements of W are also the elements of V, therefore vector addition will be associative as well as commutative in W. Thus W is an abelian group with respect to vector addition.

Now taking $\beta = 0$, we see that if $a, b \in F$ and $\alpha \in W$, then

$a\alpha + b0 \in W$ *i.e*, $a\alpha + 0 \in W$ *i.e.*, $a\alpha \in W$.

Thus W is closed under scalar multiplication.

The remaining postulates of a vector space will hold in W since they hold in V of which W is a subset. Hence $W\,(F)$ is a subspace of $V\,(F)$.

Theorem 4. *A non-empty subset W of a vector space V (F) is a subspace of V if and only if for each pair of vectors* α, β *in W and each scalar a in F the vector* $a\alpha + \beta$ *is again in W.*

Proof **The condition is necessary.** If W is a subspace of V, then W must be closed with respect to scalar multiplication and as well as with respect to vector addition. Therefore

$a \in F, \alpha \in W \Rightarrow a\alpha \in W.$

Further $a\alpha \in W, \beta \in W \Rightarrow a\alpha+\beta \in W.$

Hence the condition is necessary.

The condition is sufficient. It is given that W is a non-empty subset of V and $a \in F$, α, $\beta \in W \Rightarrow a\alpha+\beta \in W$. We are to prove that W is a subspace of V.

(i) Since W is non-empty, therefore there is at least one vector in W, say γ. Now $1 \in F \Rightarrow -1 \in F$. Therefore taking $a=-1$, $=\gamma$, $\beta=\gamma$, we get from the given condition that

$(-1)\gamma+\gamma=-(1\gamma)+\gamma=-\gamma+\gamma=\mathbf{0}$ is in W.

(ii) Now let $a \in F$, $\alpha \in W$. Since $\mathbf{0}$ is in W, therefore taking $\beta=\mathbf{0}$ in the given condition, we get $a\alpha+\mathbf{0}=a\alpha$ is in W.

Thus W is closed with respect to scalar multiplication.

(iii) Let $\alpha \in W$. Since $1 \in F$ and W is closed with respect to scalar multiplication, therefore $(-1)\,\alpha=-(1\alpha)=-\alpha$ is in W.

(iv) We have $1 \in F$. If $\alpha, \beta \in W$, then $1\alpha+\beta=\alpha+\beta$ is in W. Thus W is closed with respect to vector addition.

The remaining postulates of a vector space will hold in W, since they hold in V of which W is a subset. Hence W is a subspace of V.

Note. If we are to prove that a subset W of a vector space V is a subspace of V, then either it is sufficient to prove that

$a, b \in F$ and $\alpha, \beta \in W \Rightarrow a\alpha+b\beta \in W$

or it is sufficient to prove that

$a \in F$, and $\alpha, \beta \in W \Rightarrow a\alpha+\beta \in W.$

Examples

Example . Let $V(F)$ be any vector space. Then V itself and the subset of V consisting of zero vector only are always subspaces of V. These two are called improper subspaces. If V has any other subspace, then it is called a proper subspace. The subspace of V consisting of zero vector only is called the *zero subspace.*

Example 2. *The set W of ordered triads $(a_1, a_2, 0)$ where a_1, $a_2 \in F$ is a subspace of $V_3(F)$.* **(Meerut 1973)**

Solution. Let $\alpha=(a_1, a_2, 0)$ and $\beta=(b_1, b_2, 0)$ be any two elements of W. Then $a_1, a_2, b_1, b_2 \in F$. If a, b be any two elements of F, we have

$a\alpha+b\beta=a(a_1, a_2, 0)+b(b_1, b_2, 0)=(aa_1, aa_2, 0)+(bb_1, bb_2, 0)$
$=(aa_1+bb_1, aa_2+bb_2, 0) \in W$ since $aa_1+bb_1, aa_2+bb_2 \in F$
and the last co-ordinate of this triad is zero.

Hence W is a subspace of $V_3(F)$.

Example 3. *Let V be the vector space of all polynomials in an indeterminate x over a field F. Let W be a subset of V consisting of all polynomials of degree $\leqslant n$. Then W is a subspace of V.*

Solution. Let α and β be any two elements of W. Then α, β are polynomials over F of degree $\leqslant n$. If a, b are any two elements of F, then $a\alpha+b\beta$ will also be a polynomial of degree $\leqslant n$. Therefore $a\alpha+b\beta \in W$. Hence W is a subspace of V.

Example 4. *If a_1, a_2, a_3 are fixed elements of a field F, then the set W of all ordered triads (x_1, x_2, x_3) of elements of F, such that*

$$a_1x_1+a_2x_2+a_3x_3=0,$$

is a subspace of V_3 (F).

Solution. Let $\alpha=(x_1, x_2, x_3)$ and $\beta=(y_1, y_2, y_3)$ be any two elements of W. Then $x_1, x_2, x_3, y_1, y_2, y_3$ are elements of F and are such that

$$a_1x_1+a_2x_2+a_3x_3=0 \qquad \text{...(1)}$$

$$a_1y_1+a_2y_2+a_3y_3=0. \qquad \text{...(2)}$$

If a, b be any two elements of F, we have

$$a\alpha+b\beta=a\ (x_1, x_2, x_3)+b\ (y_1, y_2, y_3)$$

$$=(ax_1, ax_2, ax_3)+(by_1, by_2, by_3)=(ax_1+by_1, ax_2+by_2, ax_3+by_3).$$

Now $a_1\ (ax_1+by_1)+a_2\ (ax_2+by_2)+a_3\ (ax_3+by_3)$

$$=a\ (a_1x_1+a_2x_2+a_3x_3)+b\ (a_1y_1+a_2y_2+a_3y_3)$$

$$=a0+b0 \qquad \text{[by (1) and (2)]}$$

$$=0.$$

$\therefore\ a\alpha+b\beta=(ax_1+by_1, ax_2+by_2, ax_3+by_3) \in W.$

Hence W is a subspace of V_3 (F).

Example 5. *Let* **R** *be the field of real numbers. Which of the following are subspaces of $V_3(\mathbf{R})$?*

(*i*) $\{(x, 2y, 3z) : x, y, z \in \mathbf{R}\}$. **(Meerut 1972, 81)**

(*ii*) $\{(x, x, x) : x \in \mathbf{R}\}$.

(*iii*) $\{(x, y, z) : x, y, z$ *are rational numbers*$\}$?

Solution. (i) Let $W=\{(x, 2y, 3z) : x, y, z \in \mathbf{R}\}$.

Let $\alpha=(x_1, 2y_1, 3z_1)$ and $\beta=(x_2, 2y_2, 3z_2)$ be any two elements of W. Then $x_1, y_1, z_1, x_2, y_2, z_2$ are real numbers. If a, b are any two real numbers, then

$$a\alpha+b\beta=a\ (x_1, 2y_1, 3z_1)+b\ (x_2, 2y_2, 3z_2)$$

$$=(ax_1+bx_2, 2ay_1+2by_2, 3az_1+3bz_2)$$

$$=(ax_1+bx_2, 2\ [ay_1+by_2], 3\ [az_1+bz_2])$$

$\in W$ since ax_1+bx_2, ay_1+by_2, az_1+bz_2 are real numbers,

Thus $a, b \in \mathbf{R}$ and $\alpha, \beta \in W \Rightarrow a\alpha+b\beta \in W$.

$\therefore$ W is a subspace of $V_3(\mathbf{R})$.

(ii) Let $W=\{(x, x, x) : x \in \mathbf{R}\}$.

Let $\alpha=(x_1, x_1, x_1)$ and $\beta=(x_2, x_2, x_2)$ be any two elements of W. Then x_1, x_2 are real numbers. If a, b are any real numbers, then

$$a\alpha+b\beta=a\,(x_1, x_1, x_1)+b\,(x_2, x_2, x_2)$$
$$=(ax_1+bx_2, ax_1+bx_2, ax_1+bx_2)\in W, \text{ since } ax_1+bx_2\in\mathbf{R}.$$

Thus W is a subspace of V_3 (**R**).

(ii) Let $W=\{(x, y, z) : x, y, z$ are rational numbers$\}$.

Now $\alpha=(3, 4, 5)$ is an element of W. Also $a=\sqrt{7}$ is an element of **R**. But $a\alpha=\sqrt{7}\,(3, 4, 5)=(3\sqrt{7}, 4\sqrt{7}, 5\sqrt{7}) \notin W$ since $3\sqrt{7}, 4\sqrt{7}, 5\sqrt{7}$ are not rational numbers.

Therefore W is not closed under scalar multiplication. Hence W is not a subspace of V_3 (**R**).

§ 4. Algebra of subspaces.

Theorem 1. *The intersection of any two subspaces W_1 and W_2 of a vector space $V(F)$ is also a subspace of V (F).*

(Delhi 1969; Kumayun 77; Madras 77;
Madurai 78; Banaras 70; Meerut 81)

Proof. Since $\mathbf{0} \in W_1$ and W_2 both, therefore $W_1\cap W_2$ is not empty.

Let $\alpha, \beta\in W_1\cap W_2$ and $a, b\in F$.

Now $\alpha\in W_1\cap W_2 \Rightarrow \alpha\in W_1$ and $\alpha\in W_2$

and $\beta\in W_1\cap W_2 \Rightarrow \beta\in W_1$ and $\beta\in W_2$

Since W_1 is a subspace, therefore

$$a, b\in F \text{ and } \alpha, \beta\in W_1 \Rightarrow a\alpha+b\beta\in W_1.$$

Similarly $a, b\in F$ and $\alpha, \beta\in W_2 \Rightarrow a\alpha+b\beta\in W_2$.

Now $a\alpha+b\beta\in W_1$, $a\alpha+b\beta\in W_2 \Rightarrow a\alpha+b\beta\in W_1\cap W_2$.

Thus $a, b\in F$ and $\alpha, \beta\in W_1\cap W_2 \Rightarrow a\alpha+b\beta\in W_1\cap W_2$.

Hence $W_1\cap W_2$ is a subspace of $V(F)$.

Note. The union of two subspaces of $V(F)$ may not be a subspace of $V(F)$. For example if **R** be the field of real numbers, then $W_1=\{(0, 0, z) : z \in \mathbf{R}\}$ and $W_2=\{(0, y, 0) : y \in \mathbf{R}\}$ are two subspaces of V_3 (**R**). We have $(0, 0, 3)\in W_1$ and $(0, 5, 0)\in W_2$.

$\therefore$ $(0, 0, 3)$ and $(0, 5, 0)$ are both elements of $W_1\cup W_2$.

But $(0, 0, 3)+(0, 5, 0)=(0, 5, 3) \notin W_1\cup W_2$ since neither $(0, 5, 3) \in W_1$ nor $(0, 5, 3) \in W_2$. Thus $W_1 \cup W_2$ is not closed under vector addition. Hence $W_1 \cup W_2$ is not a subspace of V_3 (**R**).

Theorem 2. *The union of two subspaces is a subspace if and only if one is contained in the other.* **(Delhi 1969)**

Proof. Suppose W_1 and W_2 are two subspaces of a vector space V.

Let $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$. Then $W_1 \cup W_2 = W_2$ or W_1. But W_1, W_2 are subspaces and therefore, $W_1 \cup W_2$ is also a subspace.

Conversely, suppose $W_1 \cup W_2$ is a subspace.

To prove that $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

Let us assume that W_1 is not a subset of W_2 and W_2 is also not a subset of W_1.

Now W_1 is not a subset of $W_2 \Rightarrow \exists\, \alpha \in W_1$ and $\alpha \notin W_2$ ...(1)
and W_2 is not a subset of $W_1 \Rightarrow \exists\, \beta \in W_2$ and $\beta \notin W_1$. ...(2)

From (1) and (2), we have

$$\alpha \in W_1 \cup W_2 \text{ and } \beta \in W_1 \cup W_2.$$

Since $W_1 \cup W_2$ is a subspace, therefore

$$\alpha+\beta \text{ is also in } W_1 \cup W_2.$$

But $\alpha+\beta \in W_1 \cup W_2 \Rightarrow \alpha+\beta \in W_1$ or W_2.

Suppose $\alpha+\beta \in W_1$. Since $\alpha \in W_1$ and W_1 is a subspace, therefore $(\alpha+\beta)-\alpha=\beta$ is in W_1.

But from (2), we have $\beta \notin W_1$. Thus we get a contradiction. Again suppose that $\alpha+\beta \in W_2$. Since $\beta \in W_2$ and W_2 is a subspace, therefore $(\alpha+\beta)-\beta=\alpha$ is in W_2. But from (1), we have $\alpha \notin W_2$. Thus here also we get a contradiction. Hence either $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

Theorem 3. *Arbitrary intersection of subspaces i.e., the intersection of any family of subspaces of a vector space is a subspace.*

(Banaras 1968; Kanpur 69; Meerut 68)

Proof. Let $V(F)$ be a vector space and let $\{W_t : t \in T\}$ be any family of subspaces of V. Here T is an index set and is such that $\forall\, t \in T$, W_t is a subspace of V.

Let $$U = \bigcap_{t \in T} W_t = \{x \in V : x \in W_t \ \forall\, t \in T\}$$

be the intersection of this family of subspaces of V. Then to prove that U is also a subspace of V.

Obviously $U \neq \emptyset$, since at least the zero vector 0 of V is in $W_t \ \forall\, t \in T$.

Now let $a, b \in F$ and α, β be any two elements of $\bigcap_{t \in T} W_t$.

Then $\alpha, \beta \in W_t \; \forall \; t \in T$. Since each W_t is a subspace of V, therefore $a\alpha + b\beta \in W_t \; \forall \; t \in T$. Thus $a\alpha + b\beta \in \bigcap_{t \in T} W_t$.

Thus $a, b \in F$ and $\alpha, \beta \in \bigcap_{t \in T} W_t \Rightarrow a\alpha + b\beta \in \bigcap_{t \in T} W_t$.

Hence $\bigcap_{t \in T} W_t$ is a subspace of $V(F)$.

Smallest subspace containing any subset of $V(F)$. Let $V(F)$ be a vector space and S be any subset of V. If U is a subspace of V containing S and is itself contained in every subspace of V containing S, then U is called the **smallest subspace** of V containing S. The smallest subspace of V containing S is also called the subspace of V generated or spanned by S and we shall denote it by the symbol $\{S\}$ or by (S). It can be easily seen that the intersection of all the subspaces of $V(F)$ containing S is the subspace of $V(F)$ generated by S. If $\{S\} = V$, then we say that V is spanned by S.

Exercises

1. Show that the set W of the elements of the vector space $V_3(\mathbf{R})$ of the form $(x+2y, y, -x+3y)$ where $x, y \in \mathbf{R}$ is a subspace of $V_3(\mathbf{R})$. **(Meerut 1974)**

2. Let $V = \mathbf{R}^3$ and W be the set of all ordered triads (x, y, z) such that $x - 3y + 4z = 0$. Prove that W is a subspace of $\mathbf{R}^3$.

3. Which of the following sets of vectors $\alpha = (a_1, a_2, \ldots, a_n)$ in $\mathbf{R}^n$ are subspaces of $\mathbf{R}^n$ $(n > 3)$?
   (i) all α such that $a_1 \leqslant 0$;
   (ii) all α such that a_3 is an integer ;
   (iii) all α such that $a_2 + 4a_3 = 0$.

Ans. (i) not a subspace ; (ii) not a subspace ; (iii) subspace.

§ 5. Linear combination of vectors. Linear span of a set.

Linear combination. Definition. *Let $V(F)$ be a vector space. If $\alpha_1, \alpha_2, \ldots, \alpha_n \in V$, then any vector*

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n \text{ where } a_1, a_2, \ldots, a_n \in F$$

is called a linear combination of the vectors $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Linear span. Definition. *Let $V(F)$ be a vector space and S be any non-empty subset of V. Then the linear span of S is the set of all linear combinations of finite sets of elements of S and is denoted by $L(S)$.* Thus we have

$$L(S) = \{a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n : \alpha_1, \alpha_2, \ldots, \alpha_n$$

is any arbitrary finite subset of S and $a_1, a_2, \ldots, a_n$ is any arbitrary finite subset of F}.

Theorem 1. *The linear span $L(S)$ of any subset S of a vector space $V(F)$ is a subspace of V generated by S i.e., $L(S)=\{S\}$.*

Proof. Let α, β be any two elements of $L(S)$.

Then $\alpha=a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m$

and $\beta=b_1\beta_1+b_2\beta_2+\ldots+b_n\beta_n$

where the a's and b's are elements of F and the α's and β's are elements of S.

If a, b be any two elements of F, then

$$a\alpha+b\beta=a\,(a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m)+b\,(b_1\beta_1+b_2\beta_2+\ldots+b_n\beta_n)$$
$$=a\,(a_1\alpha_1)+a\,(a_2\alpha_2)+\ldots+a\,(a_m\alpha_m)+b\,(b_1\beta_1)+b(b_2\beta_2)+\ldots+b\,(b_n\beta_n)$$
$$=(aa_1)\,\alpha_1+(aa_2)\,\alpha_2+\ldots+(aa_m)\,\alpha_m+(bb_1)\,\beta_1+(bb_2)\,\beta_2+\ldots+(bb_n)\,\beta_n.$$

Thus $a\alpha+b\beta$ has been expressed as a linear combination of a finite set $\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n$ of the elements of S. Consequently $a\alpha+b\beta \in L(S)$.

Thus $a, b \in F$ and $\alpha, \beta \in L(S) \Rightarrow a\alpha+b\beta \in L(S)$.

Hence $L(S)$ is a subspace of $V(F)$.

Also each element of S belongs to $L(S)$ because if $\alpha_r \in S$, then $\alpha_r=1\alpha_r$ and this implies that $\alpha_r \in L(S)$. Thus $L(S)$ is a subspace of V and S is contained in $L(S)$.

Now if W is any subspace of V containing S, then each element of $L(S)$ must be in W because W is to be closed under vector addition and scalar multiplication. Therefore $L(S)$ will be contained in W.

Hence $L(S)=\{S\}$ *i.e.*, $L(S)$ is the smallest subspace of V containing S.

Note. If in any case we are to prove that $L(S)=V$, then we should prove that $V \subseteq L(S)$ because $L(S) \subseteq V$ since $L(S)$ is a subspace of V. In order to prove that $V \subseteq L(S)$, we should prove that each element of V can be expressed as a linear combination of a finite number of elements of S. Then each element of V will also be an element of $L(S)$ and we shall have $V \subseteq L(S)$.

Finally $V \subseteq L(S)$ and $L(S) \subseteq V \Rightarrow L(S)=V$.

Examples

Example 1. The subset containing a single element (1, 0, 0) of the vector space $V_3(F)$ generates the subspace which is the totality of elements of the form $(a, 0, 0)$.

Example 2. The subset {(1, 0, 0), (0, 1, 0)} of $V_3(F)$ generates the subspace which is the totality of the elements of the form $(a, b, 0)$.

Example 3. *The subset* $S=\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ *of* $V_3(F)$ *generates or spans the entire vector space* $V_3(F)$ *i.e.,* $L(S)=V$.

If (a, b, c) be any element of V, then

$$(a, b, c)=a(1, 0, 0)+b(0, 1, 0)+c(0, 0, 1).$$

Thus $(a, b, c)\in L(S)$. Hence $V\subseteq L(S)$. Also $L(S)\subseteq V$.

Hence $L(S)=V$.

§ 6. **Linear sum of two subspaces. Definition.** *Let* W_1 *and* W_2 *be two subspaces of the vector space* $V(F)$. *Then the linear sum of the subspaces* W_1 *and* W_2 *denoted by* W_1+W_2 *is the set of sums* $\alpha_1+\alpha_2$ *such that* $\alpha_1\in W_1$, $\alpha_2\in W_2$.

Thus $\quad W_1+W_2=\{\alpha_1+\alpha_2 : \alpha_1\in W_1, \alpha_2\in W_2\}$.

Theorem. *If* W_1 *and* W_2 *are subspaces of the vector space* $V(F)$, *then*

(*i*) W_1+W_2 *is a subspace of* $V(F)$.

(Meerut 1971, 79; Kanpur 70)

(*ii*) $W_1+W_2=\{W_1\cup W_2\}$ *i.e.,* $L(W_1\cup W_2)=W_1+W_2$.

Proof. (i) Let α, β be any two elements of W_1+W_2.

Then $\alpha=\alpha_1+\alpha_2$ and $\beta=\beta_1+\beta_2$ where $\alpha_1, \beta_1\in W_1$ and $\alpha_2, \beta_2\in W_2$. If $a, b\in F$, we have

$$a\alpha+b\beta=a\ (\alpha_1+\alpha_2)+b\ (\beta_1+\beta_2)=(a\alpha_1+b\beta_1)+(a\alpha_2+b\beta_2).$$

Since W_1 is a subspace of V, therefore $a, b\in F$

and $\quad \alpha_1, \beta_1\in W_1\Rightarrow a\alpha_1+b\beta_1\in W_1$.

Similarly $\quad a\alpha_2+b\beta_2\in W_2$.

Consequently $a\alpha+b\beta=(a\alpha_1+b\beta_1)+(a\alpha_2+b\beta_2)\in W_1+W_2$.

Thus $\quad a, b\in F$ and $\alpha, \beta\in W_1+W_2 \Rightarrow a\alpha+b\beta\in W_1+W_2$.

Hence W_1+W_2 is a subspace of $V(F)$.

(ii) Since W_2 contains the zero vector, therefore if $\alpha_1\in W_1$, then we can write $\alpha_1=\alpha_1+0\in W_1+W_2$. Thus $W_1\subseteq W_1+W_2$. Similarly $W_2\subseteq W_1+W_2$. Hence $W_1\cup W_2\subseteq W_1+W_2$. Therefore W_1+W_2 is a subspace of $V(F)$ containing $W_1\cup W_2$.

Now to prove that $W_1+W_2=\{W_1\cup W_2\}$ we should prove that $W_1+W_2\subseteq L(W_1\cup W_2)$ and $L(W_1\cup W_2)\subseteq W_1+W_2$.

Let $\alpha=\alpha_1+\alpha_2$ be any element of W_1+W_2. Then $\alpha_1\in W_1$ and $\alpha_2\in W_2$. Therefore $\alpha_1, \alpha_2\in W_1\cup W_2$. We can write

$$\alpha_1+\alpha_2=1\alpha_1+1\alpha_2$$

Thus $\alpha_1+\alpha_2$ is a linear combination of a finite number of elements $\alpha_1, \alpha_2\in W_1\cup W_2$.

Therefore $\alpha_1+\alpha_2 \in L(W_1 \cup W_2)$.

$\therefore \quad W_1+W_2 \subseteq L(W_1 \cup W_2)$.

Also $L(W_1 \cup W_2)$ is the smallest subspace containing $W_1 \cup W_2$ and W_1+W_2 is a subspace containing $W_1 \cup W_2$. Therefore $L(W_1 \cup W_2)$ must be contained in W_1+W_2. Consequently

$$L(W_1 \cup W_2) \subseteq W_1+W_2.$$

Hence $W_1+W_2=L(W_1 \cup W_2)=\{W_1 \cup W_2\}$.

Theorem. *If S, T are subsets of $V(F)$, then*

(i) $S \subseteq T \Rightarrow L(S) \subseteq L(T)$. *(ii)* $L(S \cup T)=L(S)+L(T)$.

(iii) $L(L(S))=L(S)$.

Proof. (i) Let $\alpha=a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n \in L(S)$, where $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a finite subset of S. Since $S \subseteq T$, therefore $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a finite subset of T. So $\alpha \in L(T)$.

Thus $\alpha \in L(S) \Rightarrow \alpha \in L(T)$.

$$L(S) \subseteq L(T).$$

(ii) Let α be any element of $L(S \cup T)$. Then

$$\alpha=a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m+b_1\beta_1+b_2\beta_2+\ldots+b_p\beta_p$$

where $\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_p\}$ is a finite subset of $S \cup T$ such that $\{\alpha_1, \alpha_2, \ldots, \alpha_m\} \subseteq S$ and $\{\beta_1, \beta_2, \ldots, \beta_p\} \subseteq T$.

Now $\quad a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m \in L(S)$

and $\quad b_1\beta_1+b_2\beta_2+\ldots+b_p\beta_p \in L(T)$.

Therefore $\alpha \in L(S)+L(T)$. Consequently $L(S \cup T) \subseteq L(S)+L(T)$.

Now let γ be any element of $L(S)+L(T)$. Then $\gamma=\beta+\delta$ where $\beta \in L(S)$ and $\delta \in L(T)$ Now β will be a linear combination of a finite number of elements of S and δ will be a linear combination of a finite number of elements of T. Therefore $\beta+\delta$ will be a linear combination of a finite number of elements of $S \cup T$. Thus $\beta+\delta \in L(S \cup T)$. Consequently $L(S)+L(T) \subseteq L(S \cup T)$.

Hence $L(S \cup T)=L(S)+L(T)$.

(iii) $L(L(S))$ is the smallest subspace of V containing $L(S)$. But $L(S)$ is a subspace of V. Therefore the smallest subspace of V containing $L(S)$ is $L(S)$ itself. Hence $L(L(S))=L(S)$.

§ 7. Linear dependence and linear independence of vectors.

Linear dependence. Definition. *Let $V(F)$ be a vector space. A finite set $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ of vectors of V is said to be linearly dependent if there exist scalars $a_1, a_2, \ldots, a_n \in F$ not all of them* 0 *(some of them may be zero) such that*

$$a_1\alpha_1+a_2\alpha_2+a_3\alpha_3+\ldots+a_n\alpha_n=0.$$

(Banaras 1970 : Meerut 75 ; Punjab 66)

Linear independence. Definition. *Let $V(F)$ be a vector space. A finite set $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ of vectors of V is said to be linearly independent if every relation of the form*

$$a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n = 0,\ a_i \in F,\ 1 \leqslant i \leqslant n$$
$$\Rightarrow a_i = 0 \text{ for each } 1 \leqslant i \leqslant n.$$

Any **infinite set** *of vectors of V is said to be linearly independent if its every finite subset is linearly independent, otherwise it is linearly dependent.* **(Banaras 1970; Punjab 69)**

Examples

Example 1. *Prove that if two vectors are linearly dependent, one of them is a scalar multiple of the other.*

Solution. Let α, β be two linearly dependent vectors of the vector space V. Then $\exists$ scalars a, b not both zero, such that

$$a\alpha + b\beta = 0.$$

If $a \neq 0$, then we get

$$a\alpha = -b\beta \Rightarrow a^{-1}(a\alpha) = a^{-1}(-b\beta)$$
$$\Rightarrow (a^{-1}a)\ \alpha = [a^{-1}(-b)]\beta \Rightarrow 1\alpha = (-a^{-1}b)\beta$$
$$\Rightarrow \alpha = -\frac{b}{a}\ \beta \Rightarrow \alpha \text{ is a scalar multiple of } \beta.$$

If $b \neq 0$, then we get

$$b\beta = -a\alpha \Rightarrow \beta = \left(-\frac{a}{b}\right)\alpha \Rightarrow \beta \text{ is a scalar multiple of } \alpha.$$

Thus one of the vectors α and β is a scalar multiple of the other.

Example 2. *In the vector space $V_n(F)$, the system of n vectors*

$$e_1 = (1, 0, 0, \ldots, 0),\ e_2 = (0, 1, 0, \ldots, 0, 0),\ \ldots,\ e_n = (0, 0, \ldots, 0, 1)$$

is linearly independent where 1 denotes the unity of the field F.

Solution. If $a_1, a_2, a_3, \ldots, a_n$ be any scalars, then

$$a_1e_1 + a_2e_2 + \ldots + a_ne_n = 0$$
$$\Rightarrow a_1(1, 0, 0, \ldots, 0) + a_2(0, 1, 0, \ldots, 0) + \ldots + a_n(0, 0, \ldots, 0, 1) = 0$$
$$\Rightarrow (a_1, a_2, \ldots, a_n) = (0, 0, \ldots, 0) \Rightarrow a_1 = 0, a_2 = 0, \ldots, a_n = 0.$$

Therefore the given set of n vectors is linearly independent.

In particular $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ *is a linearly independent subset of $V_3(F)$.*

Example 3. *If the set* $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ *of vectors of $V(F)$ is linearly independent, then none of the vectors $\alpha_1, \alpha_2, \ldots, \alpha_n$ can be zero vector.* **(Meerut 1974)**

Solution. Let α_r be equal to zero vector where $1 \leqslant r \leqslant n$. Then $0\alpha_1 + 0\alpha_2 + \ldots + a\alpha_r + 0\alpha_{r+1} + \ldots + 0\alpha_n = 0$ for any $a \neq 0$ in F.

Since $a \neq 0$, therefore from this relation we conclude that S is linearly dependent. Thus we get a contradiction because it is given that S is linearly independent. Hence none of the vectors $\alpha_1, \alpha_2, \ldots, \alpha_n$ can be zero vector. We also conclude that *a set of vectors which contains the zero vector is necessarily linearly dependent.*

Example 4. *Every superset of a linearly dependent set of vectors is linearly dependent.*

Solution. Let $S=\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a linearly dependent set of vectors. Then there exist scalars $a_1, a_2, \ldots, a_n$ not all zero such that

$$a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n=\mathbf{0}. \quad \ldots(1)$$

Now let $S'=\{\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_m\}$ be a superset of S. Then we have from (1)

$$a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n+0\beta_1+0\beta_2+\ldots+0\beta_m=\mathbf{0}. \quad \ldots(2)$$

Since in the relation (2) the scalar coefficients are not all 0, therefore S' is linearly dependent.

From this we also conclude that *any subset of a linearly independent set of vectors is also linearly independent.*

Example 5. *A system consisting of a single non-zero vector is always linearly independent.*

Solution. Let $S=\{\alpha\}$ be a subset of a vector space V and let α be not equal to zero vector. If a is any scalar, then

$$a\alpha=\mathbf{0}$$
$$\Rightarrow a=0 \qquad \text{[Since } \alpha \text{ is not zero vector]}$$

$\therefore$ the set S is linearly independent.

Example 6. *Show that*

$$S=\{(1, 2, 4), (1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

is a linearly dependent subset of the vector space V_3 (**R**) *where* **R** *is the field of real numbers.*

Solution. We have

$1\,(1, 2, 4)+(-1)\,(1, 0, 0)+(-2)\,(0, 1, 0)+(-4)\,(0, 0, 1)$
$=(1, 2, 4)+(-1, 0, 0)+(0, -2, 0)+(0, 0, -4)$
$=(0, 0, 0)$ *i.e.*, zero vector.

Since in this relation the scalar coefficients $1, -1, -2, -4$ are not all zero, therefore the given system S is linearly dependent.

Example 7. *In* V_3 (**R**), *where* **R** *is the field of real numbers, examine each of the following sets of vectors for linear dependence :*

(*i*) $\{(2, 1, 2), (8, 4, 8)\}$
(*ii*) $\{(1, 2, 0), (0, 3, 1), (-1, 0, 1)\}$
(*iii*) $\{(-1, 2, 1), (3, 0, -1), (-5, 4, 3)\}$

(*iv*) $\{(2, 3, 5), (4, 9, 25)\}$.

(*v*) $\{(1, 2, 1), (3, 1, 5), (3, -4, 7)\}$. **(Meerut 1979)**

Solution. (i) We have

$4\,(2, 1, 2)+(-1)\,(8, 4, 8)=(8, 4, 8)+(-8, -4, -8)$
$=(0, 0, 0)$ *i.e.*, the zero vector.

Since in this relation the scalar coefficients $4, -1$ are not both zero, therefore the given set is linearly dependent.

(ii) Let a, b, c be scalars *i.e.*, real numbers such that

$a\,(1, 2, 0)+b\,(0, 3, 1)+c\,(-1, 0, 1)=(0, 0, 0)$
i.e., $(a-c, 2a+3b, b+c)=(0, 0, 0)$
i.e., $a+0b-c=0,\ 2a+3b+0c=0,\ 0a+b+c=0.$

These equations will have a non-zero solution *i.e.* a solution in which a, b, c are not all zero iff the rank of the coefficient matrix is less than three *i.e.*, the number of unknowns a, b, c. If the rank is 3, then the zero solution $a=0, b=0, c=0$ will be the only solution.

Coefficient matrix $A=\begin{bmatrix}1 & 0 & -1\\ 2 & 3 & 0\\ 0 & 1 & 1\end{bmatrix}$.

We have $|A|=1\,(3-0)-2\,(0+1)=1\neq 0.$

$\therefore$ Rank $A=3$. Hence $a=0, b=0, c=0$ is the only solution. Therefore the given system is linearly independent.

(iii) Let a, b, c be scalars such that

$a\,(-1, 2, 1)+b\,(3, 0, -1)+c\,(-5, 4, 3)=(0, 0, 0)$
i.e., $(-a+3b-5c, 2a+0b+4c, a-b+3c)=(0, 0, 0)$
i.e., $-a+3b-5c=0,\ 2a+0b+4c=0,\ a-b+3c=0.$

The coefficient matrix A of these equations is

$$A=\begin{bmatrix}-1 & 3 & -5\\ 2 & 0 & 4\\ 1 & -1 & 3\end{bmatrix}.$$

We have $|A|=-1\,(0+4)-2\,(9-5)+1\,(12-0)=0.$

$\therefore$ Rank A is <3 *i.e.*, the number of unknowns a, b, c. Therefore the given system of equations will possess a non-zero solution. For example $a=-2, b=1, c=1$, is a non-zero solution. Hence the given system of vectors is linearly dependent.

(iv) Let a, b be scalars *i.e.*, real numbers such that

$a\,(2, 3, 5)+b\,(4, 9, 25)=(0, 0, 0)$
i.e., $(2a+4b, 3a+9b, 5a+25b)=(0, 0, 0)$
i.e., $2a+4b=0,\ 3a+9b=0,\ 5a+25b=0.$

The coefficient matrix A of these equations is

$$\begin{bmatrix} 2 & 4 \\ 3 & 9 \\ 5 & 25 \end{bmatrix}.$$

Obviously rank $A=2$ *i.e.*, equal to the number of unknowns a and b. Therefore these equations have the only solution $a=0$, $b=0$.

Hence the given set of vectors is linearly independent.

(v) Let a, b, c be scalars *i.e.*, real numbers such that

$a(1, 2, 1)+b(3, 1, 5)+c(3, -4, 7)=(0, 0, 0)$

i.e., $(a+3b+3c, 2a+b-4c, a+5b+7c)=(0, 0, 0)$

i.e.,
$$a+3b+3c=0, \qquad \text{...(1)}$$
$$2a+b-4c=0, \qquad \text{...(2)}$$
$$a+5b+7c=0. \qquad \text{...(3)}$$

Multiplying (1) by 2, we get
$$2a+6b+6c=0. \qquad \text{...(4)}$$

Subtracting (4) from (2), we get
$$-5b-10c=0$$
or
$$b+2c=0. \qquad \text{...(5)}$$

Again subtracting (3) from (1), we get
$$-2b-4c=0, \quad \text{or} \quad b+2c=0. \qquad \text{...(6)}$$

The equations (5) and (6) are the same and give $b=-2c$. Putting $b=-2c$ in (1), we get $a=3c$. If we take $c=1$, we get $b=-2$ and $a=3$. Thus $a=3, b=-2, c=1$ is a non-zero solution of the equations (1), (2) and (3). Hence the given set of vectors is linearly dependent.

Example 8. *If F is the field of complex numbers, prove that the vectors (a_1, a_2) and (b_1, b_2) in $V_2(F)$ are linearly dependent iff*
$$a_1b_2-a_2b_1=0.$$

Solution. Let $x, y \in F$. Then
$$x(a_1, a_2)+y(b_1, b_2)=(0, 0)$$
$$\Rightarrow (xa_1+yb_1, xa_2+yb_2)=(0, 0).$$

$$\left.\begin{aligned} \text{Therefore} \quad & a_1x+b_1y=0 \\ \text{and} \quad & a_2x+b_2y=0 \end{aligned}\right\}.$$

The necessary and sufficient condition for these equations to possess a non-zero solution is that

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}=0 \text{ } i.e., a_1b_2-a_2b_1=0.$$

Hence the given system is linearly dependent iff $a_1b_2-a_2b_1=0$.

Example 9. *If α_1 and α_2 are vectors of $V(F)$, and $a, b \in F$, show that the set $\{\alpha_1, \alpha_2, a\alpha_1+b\alpha_2\}$ is linearly dependent.*

Solution. We have

$$(-a)\,\alpha_1+(-b)\,\alpha_2+1\,(a\alpha_1+b\alpha_2)$$
$$=(-a+a)\,\alpha_1+(-b+b)\,\alpha_2=0\alpha_1+0\alpha_2=\mathbf{0}\text{ i.e., zero vector.}$$

In the above linear combination the scalar coefficient $1\neq 0$. Therefore whatever may be the scalars $-a$ and $-b$, the given set of vectors is linearly dependent.

Example 10. *Let α_1, α_2, α_3 be vectors of $V(F)$, $a, b \in F$. Show that the set $\{\alpha_1, \alpha_2, \alpha_3\}$ is linearly dependent if the set $\{\alpha_1+a\alpha_2+b\alpha_3, \alpha_2, \alpha_3\}$ is linearly dependent.* **(Meerut 1969)**

Solution. Since the set $\{\alpha_1+a\alpha_2+b\alpha_3, \alpha_2, \alpha_3\}$ is linearly dependent, therefore there exist scalars x, y, z not all zero such that

$$x\,(\alpha_1+a\alpha_2+b\alpha_3)+y\alpha_2+z\alpha_3=\mathbf{0}$$

i.e., $$x\alpha_1+(xa+y)\alpha_2+(xb+z)\,\alpha_3=\mathbf{0}. \qquad \ldots(1)$$

If in the relation (1), the coefficients x, $xa+y$, $xb+z$ are not all zero, then the set $\{\alpha_1, \alpha_2, \alpha_3\}$ will also be linearly dependent.

If $x\neq 0$, then the problem is at once solved whatever y and z may be. However if $x=0$, then at least one of y and z is not zero. Therefore at least one of $xa+y$ and $xb+z$ will not be zero since when $x=0$ then $xa+y$ and $xb+z$ reduce to y and z respectively.

Hence in the relation (1) the scalar coefficients of $\alpha_1, \alpha_2, \alpha_3$ are not all zero. Therefore the set $\{\alpha_1, \alpha_2, \alpha_3\}$ is also linearly dependent.

Example 11. *If α, β, γ are linearly independent vectors of $V(F)$ where F is the field of complex numbers, then so also are $\alpha+\beta$, $\beta+\gamma$, $\gamma+\alpha$.* **(Banaras 1970)**

Solution. Let a, b, c be scalars such that

$$a(\alpha+\beta)+b\,(\beta+\gamma)+c\,(\gamma+\alpha)=\mathbf{0}$$

i.e., $$(a+c)\,\alpha+(a+b)\,\beta+(b+c)\,\gamma=\mathbf{0}. \qquad \ldots(1)$$

But α, β, γ are linearly independent. Therefore (1) implies

$$a+0b+c=0,\ a+b+0c=0,\ 0a+b+c=0.$$

The coefficient matrix A of these equations is

$$A=\begin{bmatrix}1 & 0 & 1\\ 1 & 1 & 0\\ 0 & 1 & 1\end{bmatrix}.$$

We have rank $A=3$ *i.e.,* the number of unknowns a, b, c. Therefore $a=0$, $b=0$, $c=0$ is the only solution of the given equations. Hence $\alpha+\beta$, $\beta+\gamma$, $\gamma+\alpha$ are also linearly independent.

Example 12. *If α, β, γ are linearly independent vectors of $V(F)$ where F is the field of complex numbers, then so also are*

$$\alpha+\beta,\ \alpha-\beta,\ \alpha-2\beta+\gamma.$$

Solution Let a, b, c be scalars such that

$$a(\alpha+\beta)+b(\alpha-\beta)+c(\alpha-2\beta+\gamma)=0 \qquad \text{...(1)}$$

i.e., $$(a+b+c)\alpha+(a-b-2c)\beta+c\gamma=0. \qquad \text{...(2)}$$

But α, β, γ are linearly independent. Therefore (2) implies

$$a+b+c=0,\ a-b-2c=0,\ c=0.$$

The only solution of these equations is $c=0$, $a=0$, $b=0$.

Thus (1) implies $a=0$, $b=0$, $c=0$. Therefore the vectors $\alpha+\beta$, $\alpha-\beta$, $\alpha-2\beta+\gamma$ are linearly independent.

Example 13. *Show that the set* $\{1, x, 1+x+x^2\}$ *is a linearly independent set of vectors in the vector space of all polynomials over the real number field.* **(Meerut 1976)**

Solution. Let a, b, c be scalars (real numbers) such that

$$a(1)+bx+c(1+x+x^2)=0.$$

We have $a(1)+bx+c(1+x+x^2)=0$

$$\Rightarrow (a+c)+(b+c)x+cx^2=0$$

$$\Rightarrow a+c=0,\ b+c=0,\ c=0 \Rightarrow c=0,\ b=0,\ a=0.$$

$\therefore$ the vectors 1, x, $1+x+x^2$ are linearly independent over the field of real numbers

Example 14. *In the vector space $F[x]$ of all polynomials over the field F the infinite set $S=\{1, x, x^2, x^3, \ldots\}$ is linearly independent.*

Solution. Let $S'=\{x^{m_1}, x^{m_2}, \ldots, x^{m_n}\}$ be any finite subset of S having n vectors. Here $m_1, m_2, \ldots, m_n$ are some non-negative integers. Let $a_1, a_2, \ldots, a_n$ be scalars such that

$$a_1x^{m_1}+a_2x^{m_2}+\ldots+a_nx^{m_n}=0$$

$$(\textit{i.e.}, \text{ zero polynomial}) \qquad \text{...(1)}$$

By the definition of equality of two polynomials we have from (1) $$a_1=0,\ a_2=0, \ldots, a_n=0.$$

Thus every finite subset of S is linearly independent.

Therefore S is linearly independent.

Example 15. *Is the vector* (2, −5, 3) *in the subspace of* $\mathbf{R}^3$ *spanned by the vectors* (1, −3, 2), (2, −4, −1), (1, −5, 7)?

Solution. Let $\alpha=(2,-5,3)$, $\alpha_1=(1,-3,2)$, $\alpha_2=(2,-4,-1)$, $\alpha_3=(1,-5,7)$. If α can be expressed as a linear combination of the vectors α_1, α_2, α_3, then it will be in the subspace of $\mathbf{R}^3$ spanned by these vectors otherwise it will not be.

Let $\alpha=a_1\alpha_1+a_2\alpha_2+a_3\alpha_3$ where $a_1, a_2, a_3 \in \mathbf{R}$.

Then $(2, -5, 3)=a_1(1, -3, 2)+a_2(2, -4, -1)+a_3(1, -5, 7)$

or $(2, -5, 3)=(a_1+2a_2+a_3, -3a_1-4a_2-5a_3, 2a_1-a_2+7a_3)$.

$$\therefore \left.\begin{aligned} a_1+2a_2+a_3&=2 \\ -3a_1-4a_2-5a_3&=-5 \\ 2a_1-a_2+7a_3&=3 \end{aligned}\right\} \quad \begin{aligned} &\ldots(1)\\ &\ldots(2)\\ &\ldots(3)\end{aligned}$$

Multiplying the equation (1) by 3 and adding to (2), we get

$$2a_2-2a_3=1 \quad \text{or} \quad a_2-a_3=\tfrac{1}{2}. \qquad \ldots(4)$$

Again multiplying the equation (1) by 2 and subtracting from (3), we get

$$-5a_2+5a_3=-1 \quad \text{or} \quad a_2-a_3=1/5. \qquad \ldots(5)$$

The relations (4) and (5) show that the above equations are inconsistent. Hence the vector α cannot be expressed as a linear combination of the vectors $\alpha_1, \alpha_2, \alpha_3$. Therefore α is not in the subspace of $\mathbf{R}^3$ generated by the vectors $\alpha_1, \alpha_2, \alpha_3$.

Example 16. *In the vector space* $\mathbf{R}^3$, *let* $\alpha=(1, 2, 1)$, $\beta=(3, 1, 5)$, $\gamma=(3, -4, 7)$. *Show that the subspaces spanned by* $S=\{\alpha, \beta\}$ *and* $T=\{\alpha, \beta, \gamma\}$ *are the same.* **(Meerut 1977)**

Solution. First we shall show that the vector γ can be expressed as a linear combination of the vectors α and β. Let

$$(3, -4, 7)=a(1, 2, 1)+b(3, 1, 5).$$

Then $a+3b=3$, $2a+b=-4$, $a+5b=7$. Solving the first two equations, we get $a=-3$, $b=2$ and these satisfy the third equation also. Therefore we can write $\gamma=-3\alpha+2\beta$.

Now $\quad S \subseteq T \Rightarrow L(S) \subseteq L(T)$.

Further let $\delta \in L(T)$. Then δ can be expressed as a linear combination of the vectors α, β and γ. In this linear combination the vector γ can be replaced by $-3\alpha+2\beta$. Thus δ can be expressed as a linear combination of the vectors α and β. Therefore $\delta \in L(S)$. Thus $\delta \in L(T) \Rightarrow \delta \in L(S)$. Therefore $L(T) \subseteq L(S)$.

Hence $\quad L(T)=L(S)$.

Exercises

1. Show that the three vectors $(1, 1, -1)$, $(2, -3, 5)$ and $(-2, 1, 4)$ of $\mathbf{R}^3$ are linearly independent.

2. Show that the vectors $(1, 1, 2, 4)$, $(2, -1, -5, 2)$, $(1, -1, -4, 0)$ and $(2, 1, 1, 6)$ are linearly dependent in $\mathbf{R}^4$. **(Meerut 1971)**

3. Show that the vectors $(1, 1, 0, 0)$, $(0, 1, -1, 0)$, $(0, 0, 0, 3)$ in $\mathbf{R}^4$ are linearly independent.

4. Is the vector $(3, -1, 0, -1)$ in the subspace of $\mathbf{R}^4$ spanned by the vectors $(2, -1, 3, 2)$, $(-1, 1, 1, -3)$ and $(1, 1, 9, -5)$?

Ans. No.

5. Show that the set $\{1, x, x(1-x)\}$ is a linearly independent set of vectors in the space of all polynomials over the real number field. **(Meerut 1971)**

6. Find whether the vectors $2x^3+x^2+x+1$, x^3+3x^2+x-2 and x^3+2x^2-x+3 of $\mathbf{R}[x]$, the vector space of all polynomials over the real number field, are linearly independent or not ?

(Meerut 1975) Ans. Linearly independent.

7. Prove that a set of vectors which contains the zero vector is linearly dependent. **(Meerut 1974)**

8. Show that the system of three vectors $(1, 3, 2)$, $(1, -7, -8)$, $(2, 1, -1)$ of $V_3(\mathbf{R})$ is linearly dependent.

(Meerut 1975; Madras 77)

9. Determine whether the following set of vectors in $V_3(\mathbf{Q})$ is linearly dependent or independent, $\mathbf{Q}$ being the field of rational numbers : $\{(-1, 2, 1), (3, 1, -2)\}$. **(Meerut 1974)**

Ans. Linearly independent.

10. Prove that any finite set S of vectors, not all the zero vectors, contains a linearly independent subset T which spans the same space as S.

11. Find a linearly independent subset T of the set

$$S=\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\},$$

where
$$\alpha_1=(1, 2, -1),\ \alpha_2=(-3, -6, 3),$$
$$\alpha_3=(2, 1, 3),\ \alpha_4=(8, 7, 7) \in \mathbf{R}^3$$

which spans the same space as S. **Ans.** $T=\{\alpha_1, \alpha_3\}$.

§ 8. Some Theorems on Linear Dependence and Linear Independence.

Theorem 1. *Let $V(F)$ be a vector space. If $\alpha_1, \alpha_2, \ldots, \alpha_n$ are non-zero vectors $\in V$ then either they are linearly independent or some α_k, $2 \leqslant k \leqslant n$, is a linear combination of the preceding ones, $\alpha_1, \alpha_2, \ldots, \alpha_{k-1}$.* **(Banaras 1969; Kanpur 70)**

Proof. If $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n$ are linearly independent we are nothing to prove. So let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be linearly dependent. Then there exists a relation of the form

$$a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n=0 \qquad \ldots(1)$$

where, not all the scalar coefficients $a_1, a_2, \dots, a_n$ are 0. Let k be the largest integer for which $a_k \neq 0$ *i.e.*, $a_{k+1}=0$, $a_{k+2}=0, \dots, a_n=0$ and $a_k \neq 0$. There is no harm in this assumption because at the most if $a_n \neq 0$ then $k=n$.

Also $2 \leqslant k$. Because if $a_2=0$, $a_3=0, \dots, a_n=0$, then $a_1\alpha_1=0$ and $\alpha_1 \neq 0 \Rightarrow a_1=0$. This contradicts the fact that not all the a's are 0.

Now the relation (1) reduces to

$$a_1\alpha_1+a_2\alpha_2+\dots+a_k\alpha_k=0, \text{ where } a_k \neq 0$$

or $$a_k\alpha_k=-a_1\alpha_1-a_2\alpha_2-\dots-a_{k-1}\alpha_{k-1}$$

or $$a_k^{-1}(a_k\alpha_k)=a_k^{-1}(-a_1\alpha_1-a_2\alpha_2-\dots-a_{k-1}\alpha_{k-1})$$

or $$\alpha_k=(-a_k^{-1}a_1)\alpha_1+(-a_k^{-1}a_2)\alpha_2+\dots+(-a_k^{-1}a_{k-1})\alpha_{k-1}.$$

Thus α_k is a linear combination of its preceding vectors.

Theorem 2. *The set of non-zero vectors $\alpha_1, \alpha_2, \dots, \alpha_n$ of $V(F)$ is linearly dependent if some α_k, $2 \leqslant k \leqslant n$, is a linear combination of the preceding ones.* **(Banaras 1969; Kanpur 70)**

Proof. If some α_k, $2 \leqslant k \leqslant n$ is a linear combination of the preceding ones, $\alpha_1, \alpha_2, \dots, \alpha_{k-1}$, then $\exists$ scalars $a_1, a_2, \dots, a_{k-1}$ such that $\alpha_k=a_1\alpha_1+a_2\alpha_2+\dots+a_{k-1}\alpha_{k-1}$

$$\Rightarrow 1\alpha_k-a_1\alpha_1-a_2\alpha_2-\dots-a_{k-1}\alpha_{k-1}=0 \qquad \dots(1)$$

$\Rightarrow$ the set $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ is linearly dependent because in the linear combination (1) the scalar coefficient $1 \neq 0$.
Hence the set $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ of which $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ is a subset must be linearly dependent.

Theorem 3. *If in a vector space $V(F)$, a vector β is a linear combination of the set of vectors $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n$, then the set of vectors $\beta, \alpha_1, \alpha_2, \dots, \alpha_n$ is linearly dependent.*

Proof. Since β is a linear combination of $\alpha_1, \alpha_2, \dots, \alpha_n$ therefore there exist scalars $a_1, a_2, \dots, a_n$ such that

$$\beta=a_1\alpha_1+a_2\alpha_2+\dots+a_n\alpha_n$$

$$\Rightarrow 1\beta-a_1\alpha_1-a_2\alpha_2-\dots-a_n\alpha_n=0. \qquad \dots(1)$$

In the relation (1) the scalar coefficient of β is 1 which is $\neq 0$. Hence in the relation (1) not all the scalar coefficients are 0. Therefore the set of vectors $\beta, \alpha_1, \alpha_2, \dots, \alpha_n$ is linearly dependent.

Theorem 4. *The set of non-zero vectors $\alpha_1, \alpha_2, \dots, \alpha_n$ of $V(F)$ is linearly dependent iff one of these vectors is a linear combination of the remaining $(n-1)$ vectors.* **(Kolhapur 1973; Meerut 69)**

Proof. This theorem can be easily proved.

§ 9. Basis of a Vector Space. Definition.

(I A.S. 1974; Poona 73; Madras 78; Meerut 66; Kanpur 69; Banaras 63; Punjab 69)

A sub-set S of a vector space $V(F)$ is said to be a basis of $V(F)$, if

(*i*) *S consists of linearly independent vectors.*

(*ii*) *S generates $V(F)$ i.e., $L(S)=V$ i.e., each vector in V is a linear combination of a finite number of elements of S.*

Example 1. *A system S consisting of n vectors*

$e_1=(1, 0, 0,\ldots, 0), e_2=(0, 1, 0,\ldots, 0),\ldots, e_n=(0, 0,\ldots, 0, 1)$ *is a basis of $V_n(F)$.*

Solution. Fisrt we should show that S is a linearly independent set of vectors. We have proved it in one of the previous examples.

Now we should prove that $L(S)=V_n(F)$. We have always $L(S) \subseteq V_n(F)$. So we should prove that $V_n(F) \subseteq L(S)$ *i.e.*, each vector in $V_n(F)$ is a linear combination of elements of S.

Let $\alpha=(a_1, a_2,\ldots, a_n)$ be any vector in $V_n(F)$. We can write

$$(a_1, a_2,\ldots, a_n)=a_1(1, 0,\ldots, 0)+a_2(0, 1, 0,\ldots, 0)+\ldots$$
$$+a_n(0, 0,\ldots, 0, 1)$$

i.e., $\alpha=a_1e_1+a_2e_2+\ldots+a_ne_n$.

Hence S is a basis of $V_n(F)$. We shall call this particular basis the **standard basis** of $V_n(F)$.

Note. The set $\{(1, 0), (0, 1)\}$ is a basis of $V_2(F)$. The set $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ is a basis of $V_3(F)$. As a particular case a basis of $F(F)$ is the set consisting of only the unit element of F.

Example 2. *Show that the infinite set*

$$S=\{1, x, x^2,\ldots, x^n,\ldots\}$$

is a basis of the vector space $F[x]$ of polynomials over the field F.

Solution. First we should prove that S is a linearly independent set of vectors. For proof refer some previous example.

Now we should show that S spans $F[x]$ *i.e.*, each polynomial in $F[x]$ can be expressed as a linear combination of a finite number of elements of S.

Let $f(x)=a_0+a_1x+a_2x^2+\ldots+a_tx^t$ be a polynomial of degree t.

Then $f(x)=(a_0)\,1+a_1x+a_2x^2+\ldots+a_tx^t$.

Hence S is a basis of $F[x]$.

Note. The vector space $F[x]$ has no finite basis. If we take any finite set W of polynomials, we can find a polynomial of degree greater than that of each of them. Such a polynomial cannot at any cost be expressed as a linear combination of the elements of W.

§ 10. Finite Dimensional Vector Spaces. Definition. *The vector space $V(F)$ is said to be* **finite dimensional** *or* **finitely generated** *if there exists a* **finite** *subset S of V such that $V=L(S)$.*

The vector space $V_n(F)$ of n-tuples is a finite dimensional vector space.

The vector space $F[x]$ of all polynomials over a field F is not finite dimensional. There exists no finite subset S of $F[x]$ which spans $F[x]$. *A vector space which is not finitely generated may be referred to as an infinite dimensional space.* Thus the vector space $F[x]$ of all polynomials over a field F is infinite dimensional.

Existence of basis of a finite dimensional vector space.

Theorem. *There exists a basis for each finite dimensional vector space.* **(Meerut 1980; Banaras 69; Punjab 69; Aligarh 67)**

Proof. Let $V(F)$ be a finitely generated vector space. Let $S=\{\alpha_1, \alpha_2, \dots, \alpha_m\}$ be a finite subset of V such that $L(S)=V$. Without loss of generality we may suppose that no member of S is $\mathbf{0}$.

If S is linearly independent, then S itself is a basis of V.

If S is linearly dependent, then there exists, a vector $\alpha_i \in S$ which can be expressed as a linear combination of the preceding vectors $\alpha_1, \alpha_2, \dots, \alpha_{i-1}$.

If we omit this vector α_i from S, then the remaining set S' of $m-1$ vectors

$$\alpha_1, \alpha_2, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_m$$

also generates V *i.e.*, $V=L(S')$. For if α is any element of V, then $L(S)=V$ implies that α can be written as a linear combination of $\alpha_1, \alpha_2, \dots, \alpha_m$. In this linear combination we can replace α_i by a linear combination of $\alpha_1, \alpha_2, \dots, \alpha_{i-1}$. Then α will be a linear combination of $\alpha_1, \alpha_2, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_m$. Thus V will be equal to $L(S')$.

If S' is linearly independent, the S' will be a basis of V. If S' is linearly dependent, then proceeding as above we shall get a new set of $m-2$ vectors which generates V. Continuing this process we shall, after a finite number of steps, obtain a linearly independent subset of S which generates V and which is therefore a basis of V.

At the most it may happen that we shall be left with a subset of S which contains only one non-zero vector and which spans V. We know that a set containing a single non-zero vector is definitely linearly independent and so it will form a basis of V.

Note. The above theorem may also be stated as below :

If a finite set S of vectors spans a finite dimensional vector space $V(F)$, there exists a subset of S which forms a basis of V.

(Madras 1974)

Invariance of number of elements in the basis of a finite dimensional vector space.

Dimension Theorem for Vector Spaces. *If $V(F)$ is a finite dimensional vector space, then any two bases of V have the same number of elements.*

[Punjab 1969; I.A.S 75; Kanpur ’0; Indore 70; Vikram 76; Banaras 69; Madras 78; Marathwada 72; Meerut 80]

Proof. Suppose $V(F)$ is a finite dimensional vector space. Then V definitely possesses a basis. Let $S_1=\{\alpha_1, \alpha_2, ..., \alpha_m\}$ and $S_2=\{\beta_1, \beta_2, ..., \beta_n\}$ be two bases of V. We shall prove that $m=n$.

Since $V=L(S_1)$ and $\beta_1 \in V$, therefore β_1 can be expressed as a linear combination of $\alpha_1, \alpha_2, ..., \alpha_m$. Consequently the set $S_3=\{\beta_1, \alpha_1, \alpha_2, ..., \alpha_m\}$ which also obviously generates $V(F)$ is linearly dependent. Therefore there exists a member $\alpha_i \neq \beta_1$ of this set S_3 such that α_i is a linear combination of the preceding vectors $\beta_1, \alpha_1, \alpha_2, ..., \alpha_{i-1}$. If we omit the vector α_i from S_3 then V is also generated by the remaining set

$$S_4=\{\beta_1, \alpha_1, \alpha_2, ..., \alpha_{i-1}, \alpha_{i+1}, ..., \alpha_m\}.$$

Since $V=L(S_4)$ and $\beta_2 \in V$, therefore β_2 can be expressed as a linear combination of the vectors belonging to S_4. Consequently the set

$$S_5=\{\beta_2, \beta_1, \alpha_1, \alpha_2, ..., \alpha_{i-1}, \alpha_{i+1}, ..., \alpha_m\}$$

is linearly dependent. Therefore there exists a member α_j of this set S_5 such that α_j is a linear combination of the preceding vectors. Obviously α_j will be different from β_1 and β_2 since $\{\beta_1, \beta_2\}$ is a linearly independent set. If we exclude the vector α_j from S_5, then the remaining set will generate $V(F)$.

We may continue to proceed in this manner. Here each step consists in the exclusion of an α and the inclusion of a β in the set S_1.

Obviously the set S_1 of α's cannot be exhausted before the set S_2 of β's, otherwise $V(F)$ will be a linear span of a proper subset of S_2 and thus S_2 will become linearly dependent. Therefore we must have $m \nless n$.

Interchanging the roles of S_1 and S_2 we shall get that $n \nless m$.

Hence $m=n$.

Example. For the vector space $V_3(F)$, both the sets

$$S_1=\{(1,0,0),(0,1,0),(0,0,1)\}$$

and

$$S_2=\{(1,0,0),(1,1,0),(1,1,1)\}$$

are bases as can be easily seen. Both these bases contain the same number of elements *i.e.*, 3.

Dimension of a finitely generated vector space. Definition. *The number of elements in any basis of a finite dimensional vector space $V(F)$ is called the dimension of the vector space $V(F)$ and will be denoted by dim V.* **(I A.S. 1974; Madras 78; Poona 73)**

The vector space $V_n(F)$ is of dimension n. The vector space $V_3(F)$ is of dimension 3. If a field F is regarded as a vector space over F, then F will be of dimension 1 and the set $S=\{1\}$ consisting of unity element of F alone is a basis of F. In fact every non-zero element of F will form a basis of F.

§ 11. Some properties of finite dimensional vector spaces.

Theorem 1. (Extension Theorem). *Every inearly independent subset of a finitely generated vector space $V(F)$ forms a part of a basis of V.*

Or

Every linearly independent subset of a finitely generated vector space $V(F)$ is either a basis of V or can be extended to form a basis of V. **(Nagpur 1970; Madras 78; Delhi Hons. 68; Kanpur 69; Banaras 67; Guru Nanak 82)**

Proof. Let $S=\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ be a linearly independent subset of a finite dimensional vector space $V(F)$. If dim $V=n$, then V has a finite basis say, $\{\beta_1, \beta_2, \ldots, \beta_n\}$. Consider the set

$$S_1=\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n\}.$$

Obviously $L(S_1)=V$. Since the α's can be expressed as linear combinations of the β's therefore the set S_1 is linearly dependent.

Therefore there is some vector of S_1 which is a linear combination of its preceding vectors. This vector cannot be any of the

α's since the α's are linearly independent. Therefore this vector must be some β, say, β_i. Now omit the vector β_i from S_1 and consider the set

$$S_2=\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_{i-1}, \beta_{i+1}, \ldots, \beta_n\}.$$

Obviously $L(S_2)=V$. If S_2 is linearly independent, then S_2 will be a basis of V and it is the required extended set which is a basis of V. If S_2 is not linearly independent, then repeating the above process a finite number of times, we shall get a linearly independent set containing $\alpha_1, \alpha_2, \ldots, \alpha_m$ and spanning V. This set will be a basis of V and it will contain S. Since each basis of V contains the same number of elements, therefore exactly $n-m$ elements of the set of β's will be adjoined to S so as to form a basis of V.

Theorem 2. *Each set of $(n+1)$ or more vectors of a finite dimensional vector space $V(F)$ of dimension n is linearly dependent.*

(Meerut 1970)

Proof. Let $V(F)$ be a finite dimensional vector space of dimension n. Let S be a linearly independent subset of V containing $(n+1)$ or more vectors. Then S can be extended to form a basis of V. Thus we shall get a basis of V containing more than n vectors. But every basis of V will contain exactly n vectors. Hence our assumption is wrong. Therefore if S contains $(n+1)$ or more vectors, then S must be linearly dependent. From this theorem we conclude that if S contains m vectors and S is linearly independent then

$$m \leqslant n.$$

Theorem 3. *If $V(F)$ is a finite dimensional vector space of dimension n, then any set of n linearly independent vectors in V forms a basis of V.* **(Banaras 1969)**

Proof. Let $S=\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a linearly independent subset of a finite dimensional vector space $V(F)$ of dimension n. If S is not a basis of V, then it can be extended to form a basis of V. Thus we shall get a basis of V containing more than n vectors. But every basis of V must contain exactly n vectors. Therefore our assumption is wrong and S must be a basis of V.

Theorem 4. *If a set S of n vectors of a finite dimensional vector space $V(F)$ of dimension n generates $V(F)$, then S is a basis of V.*

(Banaras 1968)

Proof. Let $V(F)$ be a finite dimensional vector space of dimension n. Let $S=\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a subset of V such that $L(S)=V$.

If S is linearly independent, then S will form a basis of V. If S is not linearly independent, then there will exist a proper subset of S which will form a basis of V. Thus we shall get a basis of V containing less than n elements. But every basis of V must contain exactly n elements. Hence S cannot be linearly dependent and so S must be a basis of V.

Note. *If V is a finite dimensional vector space of dimension n, then V cannot be generated by fewer than n vectors.*

Theorem 5. Dimension of a subspace.

Each subspace W of a finite dimensional vector space $V(F)$ of dimension n is a finite dimensional space with dim $m \leqslant n$.

Also $V=W$ iff dim $V=$dim W.

(Delhi Hons. 1970; Banaras 67; Meerut 69, 71)

Proof. Let $V(F)$ be a finite dimensional vector space of dim n. Let W be a subspace of V. Any subset of W containing $(n+1)$ or more vectors is also a subset of V and any $(n+1)$ vectors in V are linearly dependent. Therefore any linearly independent set of vectors in W can contain, at the most n vectors. Let

$$S=\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$$

be a linearly independent subset of W with a maximal number of elements. We claim that S is a basis of W. The proof is as follows:

(i) By assumption S is a linearly independent subset of W.

(ii) $L(S)=W$ as we shall just show.

Let α be any element of W. Then the $(m+1)$ vectors $\alpha, \alpha_1, \alpha_2, \ldots, \alpha_m$ belonging to W are linearly dependent because we have supposed that the largest independent subset of W contains m vectors.

Now $\{\alpha_1, \alpha_2, \ldots, \alpha_m, \alpha\}$ is a linearly dependent set. Therefore there exists a vector belonging to it which can be expressed as a linear combination of the preceding vectors. Since $\alpha_1, \alpha_2, \ldots, \alpha_m$ are linearly independent, therefore this vector cannot be any of these m vectors. So it must be α itself. Thus α can be expressed as a linear combination of $\alpha_1, \alpha_2, \ldots, \alpha_m$. Hence $L(S)=W$.

$\therefore$ S is a basis of W.

$\therefore$ dim $W=m$ and $m \leqslant n$.

Now if $V=W$, then every basis of V is also a basis of W.

Hence dim $V=$ dim $W=n$.

Conversely let dim $W=$ dim $V=n$. Then to prove that $W=V$.

Let S be a basis of W. Then $L(S)=W$ and S contains n vectors. Since S is also a subset of V and S contains n linearly independent vectors, therefore S will also be a basis of V. Therefore $L(S)=V$. Hence $W=V$.

Theorem 6. *Let $S=\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of a finite dimensional vector space $V(F)$ of dimension n. Then every element α of V can be* **uniquely** *expressed as*

$\alpha=a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n$ *where* $a_1, a_2, \ldots, a_n \in F$. **(Meerut 1971)**

Proof. Since S is a basis of V, therefore $L(S)=V$. Therefore any vector $\alpha \in V$ can be expressed as $\alpha=a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n$.

To show uniqueness, let us suppose that

$$\alpha=b_1\alpha_1+b_2\alpha_2+\ldots+b_n\alpha_n.$$

Then we must show that $a_1=b_1, a_2=b_2, \ldots, a_n=b_n$.

We have $a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n=b_1\alpha_1+b_2\alpha_2+\ldots+b_n\alpha_n$

$$\Rightarrow (a_1-b_1)\,\alpha_1+(a_2-b_2)\,\alpha_2+\ldots+(a_n-b_n)\,\alpha_n=0$$

$$\Rightarrow a_1-b_1=0, a_2-b_2=0, \ldots, a_n-b_n=0 \text{ since } \alpha_1, \alpha_2, \ldots, \alpha_n \text{ are linearly independent}$$

$$\Rightarrow a_1=b_1, a_2=b_2, \ldots, a_n=b_n.$$

Hence the theorem.

Theorem 7. *If W_1, W_2 are two subspaces of a finite dimensional vector space $V(F)$, then*

dim $(W_1+W_2)=$ *dim* W_1+ *dim* W_2- *dim* $(W_1 \cap W_2)$.

(Rajasthan 1977; Kanpur 69; Meerut 70; Vikram 76)

Proof. Let dim $(W_1 \cap W_2)=k$ and let the set

$$S=\{\gamma_1, \gamma_2, \gamma_3, \ldots, \gamma_k\}$$

be a basis of $W_1 \cap W_2$. Then $S \subseteq W_1$ and $S \subseteq W_2$.

Since S is linearly independent and $S \subseteq W_1$, therefore S can be extended to form a basis of W_1. Let

$$\{\gamma_1, \gamma_2, \ldots, \gamma_k, \alpha_1, \alpha_2, \ldots, \alpha_m\}$$

be a basis of W_1. Then dim $W_1=k+m$. Similarly let

$$\{\gamma_1, \gamma_2, \ldots, \gamma_k, \beta_1, \beta_2, \ldots, \beta_t\}$$

be a basis of W_2. Then dim $W_2=k+t$.

$\therefore$ dim W_1+ dim W_2- dim $(W_1 \cap W_2)=(m+k)+(k+t)-k$
$=k+m+t$.

$\therefore$ to prove the theorem we must show that

$$\dim (W_1+W_2)=k+m+t.$$

We claim that $S_1=\{\gamma_1, \gamma_2, \ldots, \gamma_k, \alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_t\}$ is a basis of W_1+W_2.

First we show that S_1 is linearly independent. Let

$$c_1\gamma_1+c_2\gamma_2+\ldots+c_k\gamma_k+a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m+b_1\beta_1+b_2\beta_2 +\ldots+b_t\beta_t=0 \quad \ldots(1)$$

$$\Rightarrow b_1\beta_1+b_2\beta_2+\ldots+b_t\beta_t=-(c_1\gamma_1+\ldots+c_k\gamma_k+a_1\alpha_1+\ldots+a_m\alpha_m) \quad \ldots(2)$$

Now $-(c_1\gamma_1+\ldots+c_k\gamma_k+a_1\alpha_1+\ldots+a_m\alpha_m) \in W_1$ since it is a linear combination of a basis of W_1. Again

$$b_1\beta_1+b_2\beta_2+\ldots+b_t\beta_t \in W_2$$

since it is a linear combination of elements belonging to a basis of W_2.

Also by virtue of the equality (2), $b_1\beta_1+\ldots+b_t\beta_t \in W_1$. Therefore $b_1\beta_1+b_2\beta_2+\ldots+b_t\beta_t \in W_1 \cap W_2$. Therefore it can be expressed as a linear combination of the basis S of $W_1 \cap W_2$. Thus we have a relation of the form

$$b_1\beta_1+b_2\beta_2+\ldots+b_t\beta_t=d_1\gamma_1+d_2\gamma_2+\ldots+d_k\gamma_k$$

$$\Rightarrow b_1\beta_1+b_2\beta_2+\ldots+b_t\beta_t-d_1\gamma_1-d_2\gamma_2-\ldots-d_k\gamma_k=0.$$

But $\beta_1, \beta_2, \ldots, \beta_t, \gamma_1, \ldots, \gamma_k$ are linearly independent vectors. Therefore we must have $b_1=0, b_2=0, \ldots, b_t=0$.

Putting these values of b's in (1), it reduces to

$$c_1\gamma_1+c_2\gamma_2+\ldots+c_k\gamma_k+a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m=0$$

$$\Rightarrow c_1=0, c_2=0, \ldots, c_k=0, a_1=0, a_2=0, \ldots, a_m=0$$

since the vectors $\gamma_1, \gamma_2, \ldots, \gamma_k, \alpha_1, \alpha_2, \ldots, \alpha_m$ are linearly independent.

Thus the relation (1) implies that

$$c_1=0, c_2=0, \ldots, c_k=0, a_1=0, \ldots, a_m=0, b_1=0, \ldots, b_t=0.$$

Therefore the set S_1 of vectors $\gamma_1, \ldots, \gamma_k, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_t$ is linearly independent.

Now to show that $L(S_1)=W_1+W_2$.

Since W_1+W_2 is a subspace of V and each element of S_1 belongs to W_1+W_2, therefore $L(S_1) \subseteq W_1+W_2$.

Again let α be any element of W_1+W_2. Then

α = some element of W_1 + some element of W_2

= a linear combination of a basis of W_1 + a linear combination of a basis of W_2

= a linear combination of S_1.

$\therefore \quad \alpha \in L(S_1)$. Hence $W_1+W_2 \subseteq L(S_1)$.

$\therefore \quad L(S_1) = W_1+W_2$.

$\therefore$ S_1 is a basis of W_1+W_2 and consequently
$\dim(W_1+W_2)=k+m+t$.

Hence the theorem.

Solved Examples

Ex. 1. *Show that the vectors* (1, 2, 1), (2, 1, 0), (1, −1, 2) *form a basis of* $\mathbf{R}^3$. **(Kanpur 1980)**

Solution. We know that the set {(1, 0, 0), (0, 1, 0), (0, 0, 1)} forms a basis for $\mathbf{R}^3$. Therefore dim $\mathbf{R}^3=3$. If we show that the set $S=\{(1, 2, 1), (2, 1, 0), (1, -1, 2)\}$ is linearly independent, then this set will also form a basis for $\mathbf{R}^3$.

[See theorem 3, page 427]

We have

$$a_1(1, 2, 1)+a_2(2, 1, 0)+a_3(1, -1, 2)=(0, 0, 0)$$
$$\Rightarrow (a_1+2a_2+a_3,\ 2a_1+a_2-a_3,\ a_1+2a_3)=(0, 0, 0).$$

$$\therefore\quad a_1+2a_2+a_3=0 \qquad \ldots(1)$$
$$2a_1+a_2-a_3=0 \qquad \ldots(2)$$
$$a_1+2a_3=0. \qquad \ldots(3)$$

Now we shall solve these equations to get the values of a_1, a_2, a_3. Multiplying the equation (2) by 2, we get

$$4a_1+2a_2-2a_3=0. \qquad \ldots(4)$$

Subtracting (4) from (1), we get

$$-3a_1+3a_3=0$$

or

$$-a_1+a_3=0. \qquad \ldots(5)$$

Adding (3) and (5), we get $3a_3=0$ or $a_3=0$. Putting $a_3=0$ in (3), we get $a_1=0$. Now putting $a_3=0$ and $a_1=0$ in (1), we get $a_2=0$.

Thus solving the equations (1), (2) and (3), we get $a_1=0$, $a_2=0$, $a_3=0$. Therefore the set S is linearly independent. Hence it forms a basis for $\mathbf{R}^3$.

Ex. 2. *Determine whether or not the following vectors form a basis of* $\mathbf{R}^3$:

(1, 1, 2), (1, 2, 5), (5, 3, 4). **(Meerut 1980)**

Solution. We know that dim $\mathbf{R}^3=3$. If the given set of vectors is linearly independent, it will form a basis of $\mathbf{R}^3$ otherwise not. We have

$a_1(1, 1, 2)+a_2(1, 2, 5)+a_3(5, 3, 4)=(0, 0, 0)$

$\Rightarrow (a_1+a_2+5a_3,\ a_1+2a_2+3a_3,\ 2a_1+5a_2+4a_3)=(0, 0, 0).$

$$\therefore \left.\begin{aligned} a_1+a_2+5a_3&=0\\ a_1+2a_2+3a_3&=0\\ 2a_1+5a_2+4a_3&=0\end{aligned}\right\} \qquad \begin{aligned}&\ldots(1)\\&\ldots(2)\\&\ldots(3)\end{aligned}$$

Now we shall solve these equations to get the values of a_1, a_2, a_3. Subtracting (2) from (1), we get

$$-a_2+2a_3=0. \qquad \ldots(4)$$

Multiplying (1) by 2, we get

$$2a_1+2a_2+10a_3=0. \qquad \ldots(6)$$

Subtracting (5) from (3), we get

$$3a_2-6a_3=0$$

or

$$a_2-2a_3=0. \qquad \ldots(6)$$

We see that the equations (4) and (6) are the same and give $a_2=2a_3$. Putting $a_2=2a_3$ in (1), we get $a_1=-7a_3$. If we put $a_3=1$, we get $a_2=2$ and $a_1=-7$. Thus $a_1=-7$, $a_2=2$, $a_3=1$ is a non-zero solution of the equations (1), (2) and (3). Hence the given set is linearly dependent and so it does not form a basis of $\mathbf{R}^3$.

Exercises

1. For the 3-dimensional vector space $\mathbf{R}^3$ over the field of real numbers $\mathbf{R}$, determine if the set $\{(2, -1, 0), (3, 5, 1), (1, 1, 2)\}$ is a basis. **(Nagpur 1970)**

   Ans. Yes.

2. (i) Show that the vectors $(2, 1, 4)$, $(1, -1, 2)$, $(3, 1, -2)$ form a basis for $\mathbf{R}^3$.

   (ii) Show that the set $\{(1, i, 0), (2i, 1, 1), (0, 1+i, 1-i)\}$ is a basis for V_3 (C) **(Meerut 1981)**

3. Show that a system X consisting of the vectors $\alpha_1=(1, 0, 0, 0)$, $\alpha_2=(0, 1, 0, 0)$, $\alpha_3=(0, 0, 1, 0)$ and $\alpha_4=(0, 0, 0, 1)$ is a basis set of $\mathbf{R}^4$ (R). **(Kolhapur 1973)**

4. Let V be a vector space. Let W be a subspace of V generated by the vectors $\alpha_1, \ldots, \alpha_s$. Prove that W is spanned by a linearly independent subset of $\alpha_1, \ldots, \alpha_s$. **(Dibrugarh 1967)**

5. If W is a subspace of a finite dimensional vector space V; prove that any basis of W can be extended to form a basis of V. **(Nagpur 1970)**

6. If n vectors span a vector space V containing r linearly independent vectors, then show that $n \geqslant r$. **(Indore 1970)**

7. Show that a finite subset W of a vector space $V(F)$ is linearly dependent if and only if some element of W can be expressed as a linear combination of the others. **(Kolhapur 1973)**

8. Select a basis, if any, of $\mathbf{R}^3$ (**R**) from the set $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$, where $\alpha_1=(1, -3, 2)$, $\alpha_2=(2, 4, 1)$, $\alpha_3=(3, 1, 3)$, $\alpha_4=(1, 1, 1)$. **Ans.** $\{\alpha_1, \alpha_2, \alpha_4\}$.

9. Let V be a finite dimensional vector space over a field F, and let $v_1, v_2, \ldots, v_n$ be a basis of V. If $w_1, w_2, \ldots, w_m$ in V are linearly independent over F, then prove that $m \leqslant n$. **(I.A.S. 1973)**

10. Show that the set $S=\{1, x, x^2, \ldots, x^n\}$ of $n+1$ polynomials in x is a basis of the vector space P_n (**R**), of all polynomials in x (of degree at most n) over the field of real numbers. **(I.A.S. 1977; Meerut 74)**

11. Prove that the set of all $m \times n$ matrices with entries from a field F forms a finite dimensional vector space over F, under the usual operations of matrix addition and scalar multiplication. Obtain a basis for this vector space. **(Madras 1974)**

§ 12. Homomorphism of vector spaces or linear transformations.

Definition. *Let $U(F)$ and $V(F)$ be two vector spaces. Then a mapping $f : U \rightarrow V$ is called a homomorphism or a linear transformation of U* **into** *V if* :

(*i*) $f(\alpha+\beta)=f(\alpha)+f(\beta)$, $\forall\ \alpha, \beta \in U$

and (*ii*) $f(a\alpha)=af(\alpha)\ \forall\ a \in F, \forall\ \alpha \in U$.

(Delhi 1970; Poona 73; Aligarh 65)

The conditions (i) and (ii) can be combined into a single condition $f(a\alpha+b\beta)=af(\alpha)+bf(\beta)\ \forall\ a, b \in F$ and $\forall\ \alpha, \beta \in U$.

If f is a homomorphism of U **onto** V, then V is called a homomorphic image of U.

Theorem 1. *If f is a homomorphism of $U(F)$ into $V(F)$, then*

(*i*) $f(\mathbf{0})=\mathbf{0}'$ *where* $\mathbf{0}$ *and* $\mathbf{0}'$ *are the zero vectors of U and V respectively.*

(*ii*) $f(-\alpha)=-f(\alpha)\ \forall\ \alpha \in U$.

Proof. (i) Let $\alpha \in U$. Then $f(\alpha) \in V$. Since $0'$ is the zero vector of V, therefore $f(\alpha)+0'=f(\alpha)=f(\alpha+0)=f(\alpha)+f(0)$.

Now V is an abelian group with respect to addition of vectors. Therefore $f(\alpha)+0'=f(\alpha)+f(0) \Rightarrow 0'=f(0)$, by left cancellation law.

(ii) If $\alpha \in U$, then $-\alpha \in U$. Also we have

$$0'=f(0)=f[\alpha+(-\alpha)]=f(\alpha)+f(-\alpha).$$

Now $f(\alpha)+f(-\alpha)=0' \Rightarrow f(-\alpha)=$ additive inverse of $f(\alpha)$

$$\Rightarrow f(-\alpha)=-f(\alpha).$$

Note. The students may use without any confusion the same symbol 0 to denote both the zero vectors of U and V. In the relation $f(0)=0$, the zero in the left hand side is the zero vector of U and its f-image *i.e.*, the zero on the right hand side is the zero vector of V.

Kernel of a homomorphism. Definition. *Let f be a homomorphism of a vector space $U(F)$ into a vector space $V(F)$. The kernel W of f is defined as*

$$W=\{\alpha \in U : f(\alpha)=0' \text{ where } 0' \text{ is the zero vector of } V\}.$$

Thus the kernel W of f is a subset of U consisting of those elements of U which are mapped under f onto the zero vector of V. Since $f(0)=0'$, therefore at least $0 \in W$. Thus W is not empty.

Theorem 2. *The kernel of a homomorphism is a subspace.*

(Punjab 1969)

Proof. Let $U(F)$ and $V(F)$ be two vector spaces and let f be a homomorphism of U into V.

Let $0'$ be the zero vector of V and 0 be the zero vector of U. Let W be the kernel of f *i.e.*, $W=\{\alpha \in U : f(\alpha)=0'\}$.

To prove that W is a subspace of U.

Let α, β be any two elements of W. Then $f(\alpha)=0'$ and $f(\beta)=0'$. If a, b are any two elements of F, then we have

$f(a\alpha+b\beta)=af(\alpha)+bf(\beta)'$ $\quad$ [$\because$ f is a homomorphism of U into V]

$=a0'+b0'=0'+0'=0'$.

$\therefore$ $a\alpha+b\beta \in W$.

Thus $a, b \in F$ and $\alpha, \beta \in W \Rightarrow a\alpha+b\beta \in W$.

$\therefore$ W is a subspace of U.

§ 13. Isomorphism of Vector spaces.

Definition. *Let $U(F)$ and $V(F)$ be two vector spaces. Then a mapping $f : U \to V$ is called an isomorphism of U onto V if*

(i) f *is one-one,*

(ii) f *is onto,*

(iii) $f(a\alpha+b\beta)=af(\alpha)+bf(\beta)\ \forall\ a, b \in F, \alpha, \beta \in U.$

Also then the two vector spaces U and V are said to be isomorphic and symbolically we write $U(F) \cong V(F)$.

The vector space $V(F)$ is also called the isomorphic image of the vector space $U(F)$.

If f is a homomorphism of $U(F)$ into $V(F)$, then f will become an isomorphism of U into V if f is one-one. Also in addition if f is onto V, then f will become an isomorphism of U onto V.

Isomorphism of finite dimensional vector spaces.

Theorem 1. (Isomorphism theorem for vector spaces). *Two finite dimensional vector spaces over the same field are isomorphic if and only if they are of the same dimension.*

(Banaras 1970; Kanpur 69; Meerut 77)

Proof. First suppose that $U(F)$ and $V(F)$ are two finite dimensional vector spaces each of dimension n. Then to prove that $U(F) \cong V(F)$.

Let the sets of vectors $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ and $\{\beta_1, \beta_2, \ldots, \beta_n\}$ be the bases of U and V respectively.

Any vector $\alpha \in U$ can be uniquely expressed as

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n.$$

Let $f : U \to V$ be defined by

$$f(\alpha) = a_1\beta_1 + a_2\beta_2 + \ldots + a_n\beta_n.$$

Since in the expression for α as a linear combination of $\alpha_1, \alpha_2, \ldots, \alpha_n$ the scalars $a_1, a_2, \ldots, a_n$ are unique, therefore the mapping f is well-defined *i.e.*, $f(\alpha)$ is a unique element of V.

f is one-one. We have

$f(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n) = f(b_1\alpha_1 + b_2\alpha_2 + \ldots + b_n\alpha_n)$

$\Rightarrow a_1\beta_1 + a_2\beta_2 + \ldots + a_n\beta_n = b_1\beta_1 + b_2\beta_2 + \ldots + b_n\beta_n$

$\Rightarrow (a_1 - b_1)\beta_1 + \ldots + (a_n - b_n)\beta_n = \mathbf{0}'$ (zero vector of V)

$\Rightarrow a_1 - b_1 = 0, a_2 - b_2 = 0, \ldots, a_n - b_n = 0$ because

$\beta_1, \beta_2, \ldots, \beta_n$ are linearly independe

$\Rightarrow a_1 = b_1, a_2 = b_2, \ldots, a_n = b_n$

$\Rightarrow a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n = b_1\alpha_1 + b_2\alpha_2 + \ldots + b_n\alpha_n.$

$\therefore$ f is one-one.

f is onto V. If $a_1\beta_1+a_2\beta_2+\ldots+a_n\beta_n$ is any element of V, then $\exists$ an element $a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n\in U$ such that

$$f(a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n)=a_1\beta_1+a_2\beta_2+\ldots+a_n\beta_n.$$

$\therefore$ f is onto V.

f is a linear transformation. We have

$$\begin{aligned} f\,[a\,(a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n)+b\,(b_1\alpha_1+b_2\alpha_2+\ldots+b_n\alpha_n)]\\ =f\,[(aa_1+bb_1)\,\alpha_1+(aa_2+bb_2)\,\alpha_2+\ldots+(aa_n+bb_n)\,\alpha_n]\\ =(aa_1+bb_1)\,\beta_1+(aa_2+bb_2)\,\beta_2+\ldots+(aa_n+bb_n)\,\beta_n\\ =a\,(a_1\beta_1+a_2\beta_2+\ldots+a_n\beta_n)+b\,(b_1\beta_1+b_2\beta_2+\ldots+b_n\beta_n)\\ =af\,(a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n)+bf\,(b_1\alpha_1+b_2\alpha_2+\ldots+b_n\alpha_n).\end{aligned}$$

$\therefore$ f is a linear transformation.

Hence f is an isomorphism of U onto V. $\therefore$ $U\cong V$.

Conversely, let $U(F)$ and $V(F)$ be two isomorphic finite dimensional vector spaces. Then to prove that dim $U=$dim V. Let dim $U=n$. Let $S=\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of U. If f is an isomorphism of U onto V, we shall show that $S'=\{f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)\}$ is a basis of V. Then V will also be of dimension n.

First we show that S' is linearly independent.

Let $a_1f(\alpha_1)+a_2f(\alpha_2)+\ldots+a_nf(\alpha_n)=0'$ (zero vector of V)

$\Rightarrow f(a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n)=0'$ [$\because$ f is a linear transformation]

$\Rightarrow a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n=0$ [$\because$ f is one-one and $f(0)=0'$ where 0 is zero vector of U]

$\Rightarrow a_1=0, a_2=0, \ldots, a_n=0$ since $\alpha_1, \alpha_2, \ldots, \alpha_n$ are linearly independent.

$\therefore$ S' is linearly independent.

Now to prove that $L(S')=V$. For this we shall prove that any vector $\beta\in V$ can be expressed as a linear combination of the vectors of the set S'. Since f is onto V, therefore $\beta\in V\Rightarrow$ there exists $\alpha\in U$ such that $f(\alpha)=\beta$.

Let $\alpha=c_1\alpha_1+c_2\alpha_2+\ldots+c_n\alpha_n$.

Then $\beta=f(\alpha)=f(c_1\alpha_1+c_2\alpha_2+\ldots+c_n\alpha_n)$

$=c_1f(\alpha_1)+c_2f(\alpha_2)+\ldots+c_nf(\alpha_n)$.

Thus β is a linear combination of the vectors of S'.

Hence $V=L(S')$. Therefore S' is a basis of V. Since S' contains n vectors, therefore dim $V=n$.

Note. While proving the converse, we have proved that *if f is an isomorphism of U onto V then f maps a basis of U onto a basis of V.*

Theorem 2. *Every n-dimensional vector space $V(F)$ is isomorphic to $V_n(F)$.* **(Aligarh 1965; Meerut 71; Marathwada 72)**

Proof. Let $\{\alpha_1, \alpha_2, .., \alpha_n\}$ be any basis of $V(F)$. Then every vector $\alpha \in V$ can be uniquely expressed as

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + \cdot + a_n\alpha_n, \ a_i \in F.$$

The ordered n-tuple $(a_1, a_2, \ldots, a_n) \in V_n(F)$.

Let $f : V(F) \to V_n(F)$ be defined by $f(\alpha) = (a_1, a_2, \ldots, a_n)$.

Since in the expression of α as a linear combination of $\alpha_1, \alpha_2, \ldots, \alpha_n$ the scalars $a_1, a_2, \ldots, a_n$ are unique, therefore $f(\alpha)$ is a unique element of $V_n(F)$ and thus the mapping f is well-defined.

f is one-one. Let $\alpha = a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n$ and $\beta = b_1\alpha_1 + b_2\alpha_2 + \ldots + b_n\alpha_n$ be any two elements of V. We have

$$f(\alpha) = f(\beta)$$
$$\Rightarrow f(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n) = f(b_1\alpha_1 + b_2\alpha_2 + \ldots + b_n\alpha_n)$$
$$\Rightarrow (a_1, a_2, \ldots, a_n) = (b_1, b_2, \ldots, b_n)$$
$$\Rightarrow a_1 = b_1, a_2 = b_2, \ldots, a_n = b_n \Rightarrow \alpha = \beta.$$
$$\therefore \ f \text{ is one-one.}$$

f is onto $V_n(F)$. Let $(a_1, a_2, \ldots, a_n)$ be any element of $V_n(F)$. Then there exists an element $a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n \in V(F)$ such that $f(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n) = (a_1, a_2, \ldots, a_n)$.

$$\therefore \ f \text{ is onto } V_n(F).$$

f is a linear transformation. If $a, b \in F$ and $\alpha, \beta \in V(F)$, we have $f(a\alpha + b\beta)$

$$= f[a(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n) + b(b_1\alpha_1 + b_2\alpha_2 + \ldots + b_n\alpha_n)]$$
$$= f[(aa_1 + bb_1)\alpha_1 + (aa_2 + bb_2)\alpha_2 + \ldots + (aa_n + bb_n)\alpha_n]$$
$$= (aa_1 + bb_1, aa_2 + bb_2, \ldots, aa_n + bb_n)$$
$$= (aa_1, aa_2, \ldots, aa_n) + (bb_1, bb_2, \ldots, bb_n)$$
$$= a(a_1, a_2, \ldots, a_n) + b(b_1, b_2, \ldots, b_n)$$
$$= af(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n) + bf(b_1\alpha_1 + b_2\alpha_2 + \ldots + b_n\alpha_n)$$
$$= af(\alpha) + bf(\beta).$$

$\therefore$ f is a linear transformation.

$\therefore$ f is an isomorphism of $V(F)$ onto $V_n(F)$.

Hence $V(F) \cong V_n(F)$.

Example 1. *The mapping $f : V_3(F) \to V_2(F)$ defined by*

$$f(a_1, a_2, a_3) = (a_1, a_2)$$

is a homomorphism of $V_3(F)$ onto $V_2(F)$. What is the kernel of this homomorphism ?

Solution. Let $\alpha = (a_1, a_2, a_3)$ and $\beta = (b_1, b_2, b_3)$ be any two elements of $V_3(F)$. Also let a, b be any two elements of F. We have $\quad f(a\alpha + b\beta) = f[a(a_1, a_2, a_3) + b(b_1, b_2, b_3)]$

$$= f[(aa_1 + bb_1, aa_2 + bb_2, aa_3 + bb_3)] = (aa_1 + bb_1, aa_2 + bb_2)$$

$=a\,(a_1, a_2)+b\,(b_1, b_2)=af\,(a_1, a_2, a_3)+bf\,(b_1, b_2, b_3)$
$=af\,(\alpha)+bf\,(\beta)$.

$\therefore$ f is a linear transformation.

To show that f is onto $V_2\,(F)$. Let (a_1, a_2) be any element of $V_2\,(F)$. Then $(a_1, a_2, 0)\in V_3(F)$ and we have $f\,(a_1, a_2, 0)=(a_1, a_2)$. Therefore f is onto $V_2\,(F)$.

Therefore f is a homomorphism of $V_3\,(F)$ onto $V_2\,(F)$. If W is the kernel of this homomorphism then $W=\{(0, 0, a) : a\in F\}$.

We have $\forall\ a\in F, f\,(0, 0, a)=(0, 0)=$ the zero vector of $V_2\,(F)$.

Also if $f\,(a_1, a_2, a_3)=(0, 0)$, then $\vec{f}\,(a_1, a_2, a_3)=(a_1, a_2)=(0, 0)$ implies $a_1=0$, $a_2=0$. Therefore $(a_1, a_2, a_3)\in W$.

Hence W is the kernel of f.

Example 2. *If V is a finite dimensional vector space and f is an isomorphism of V into V, prove that f must map V onto V.*

Solution. Let $V\,(F)$ be a finite dimensional vector space of dimension n. Let f be an isomorphism of V into V *i.e.*, f is a linear transformation and f is one-one. To prove that f is onto V.

Let $S=\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of V. We shall first prove that

$S'=\{f\,(\alpha_1), f\,(\alpha_2), \ldots, f\,(\alpha_n)\}$ is also a basis of V. We claim that S' is linearly independent. The proof is as follows :

Let $a_1 f\,(\alpha_1)+a_2 f\,(\alpha_2)+\ldots+a_n f\,(\alpha_n)=0$ (zero vector of V)
$\Rightarrow f\,(a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n)=0$ [$\because$ f is a linear transformation]
$\Rightarrow a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n=0$ [$\because$ f is one-one and $f\,(0)=0$]
$\Rightarrow a_1=0, a_2=0, \ldots, a_n=0$ since $\alpha_1, \alpha_2, \ldots, \alpha_n$ are linearly independent.

$\therefore$ S' is linearly independent.

Now V is of dimension n and S' is a linearly independent subset of V containing n vectors. Therefore S' must be a basis of V. Therefore each vector in V can be expressed as a linear combination of the vectors belonging to S'.

Now we shall show that f is onto V. Let α be any element of V. Then there exist scalars $c_1, c_2, \ldots, c_n$ such that

$$\alpha=c_1 f\,(\alpha_1)+c_2 f\,(\alpha_2)+\ldots+c_n f\,(\alpha_n)$$
$$=f\,(c_1\alpha_1+c_2\alpha_2+\ldots+c_n\alpha_n).$$

Now $c_1\alpha_1+c_2\alpha_2+\ldots+c_n\alpha_n\in V$ and the f-image of this element is α. Therefore f is onto V. Hence f is an isomorphism of V onto V.

Example 3. *$V(F)$ and $W(F)$ are two finite dimensional vector spaces such that dim V=dim W. If f is an isomorphism of V into W prove that f must map V onto W.*

Solution. Proceed as in example 2.

Example 4. *If V is finite dimensional and f is a homomorphism of V prove that f must be one-one and so, an isomorphism.*

(G.N.D.U. Amritsar 1982)

Solution. Let $V(F)$ be a finite dimensional vector space of dimension n. Let f be a homomorphism of V onto V *i.e.*, f is a linear transformation and f is onto V. To prove that f is one-one.

Let $S=\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of V. We shall first prove that $S'=\{f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)\}$ is also a basis of V. We claim that $L(S')=V$. The proof is as follows :

Let α be any element of V. We shall show that α can be expressed as a linear combination of $f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)$. Since f is onto V, therefore $\alpha \in V$ implies that there exists $\beta \in V$ such that $f(\beta)=\alpha$. Now β can be expressed as a linear combination of $\alpha_1, \alpha_2, \ldots, \alpha_n$. Let $\beta=a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n$.

Then $\alpha=f(\beta)=f(a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n)$

$=a_1 f(\alpha_1)+a_2 f(\alpha_2)+\ldots+a_n f(\alpha_n)$.

Thus α has been expressed as a linear combination of

$$f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n).$$

Therefore $L(S')=V$.

Since V is of dimension n and S' is a subset of V containing n vectors and $L(S')=V$, therefore S' must be a basis of V. Therefore each vector in V can be expressed as a linear combination of the vectors belonging to S' and S' is linearly independent.

Now we shall show that f is one-one. Let γ, δ be any two elements of V such that

$$\gamma=c_1\alpha_1+c_2\alpha_2+\ldots+c_n\alpha_n,\ \delta=d_1\alpha_1+d_2\alpha_2+\ldots+d_n\alpha_n.$$

We have $f(\gamma)=f(\delta)$

$\Rightarrow f(c_1\alpha_1+c_2\alpha_2+\ldots+c_n\alpha_n)=f(d_1\alpha_1+d_2\alpha_2+\ldots+d_n\alpha_n)$

$\Rightarrow c_1 f(\alpha_1)+c_2 f(\alpha_2)+\ldots+c_n f(\alpha_n)=d_1 f(\alpha_1)+d_2 f(\alpha_2)+\ldots+d_n f(\alpha_n)$

$\Rightarrow (c_1-d_1) f(\alpha_1)+(c_2-d_2) f(\alpha_2)+\ldots+(c_n-d_n) f(\alpha_n)=0$

$\Rightarrow c_1-d_1=0, c_2-d_2=0, \ldots, c_n-d_n=0$ since $f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)$ are linearly independent

$\Rightarrow c_1=d_1, c_2=d_2, \ldots, c_n=d_n \Rightarrow \gamma=\delta.$

$\therefore$ f is one-one.

$\therefore$ f is an isomorphism of V onto V.

Example 5. *If V is finite dimensional and f is a homomorphism of V into itself which is not onto prove that there is some $\alpha \neq 0$ in V such that $f(\alpha)=0$.* **(Meerut 1969)**

Solution. If f is a homomorphism of V into itself, then $f(0)=0$. Suppose there is no non-zero vector α in V such that $f(\alpha)=0$. Then f is one-one. Because

$$\begin{aligned} f(\beta)=f(\gamma) &\Rightarrow f(\beta)-f(\gamma)=0 \\ &\Rightarrow f(\beta-\gamma)=0 \qquad [\because f \text{ is linear transformation}] \\ &\Rightarrow \beta-\gamma=0 \Rightarrow \beta=\gamma. \end{aligned}$$

Now V is finite dimensional and f is a linear transformation of V into itself. Since f is one-one, therefore f must be onto V. But it is given that f is not onto. Therefore our assumption is wrong. Hence there will be a non-zero vector α in V such that

$$f(\alpha)=0.$$

Exercises

1. Define linear transformation of a vector space $V(F)$ into a vector space $W(F)$. Show that the mapping
   $$T : (a, b) \to (a+2, b+3)$$
   of $V_2(\mathbf{R})$ into itself is not a linear transformation.
2. If $f : U \to V$ is an isomorphism of the vector space U into the vector space V, then a set of vectors $f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_r)$ is linearly dependent in V if and only if the set $\alpha_1, \alpha_2, \ldots, \alpha_r$ is linearly dependent in U.
3. If $f : U \to V$ is an isomorphism of the vector space U into the vector space V, then a set of vectors $\{f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_r)\}$ is linearly independent if and only if the set $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ is linearly independent in U.
4. Let f be a linear transformation from a vector space U into a vector space V. If S is a subspace of U, prove that $f(S)$ will be a subspace of V.
5. If f is an isomorphism of a vector space V onto a vector space W, prove that f maps a basis of V onto a basis of W.

§ 14. **Quotient space.** Let W be any subspace of a vector space $V(F)$. Then W is a subgroup of the abelian group of V with respect to addition of vectors. If $\alpha \in V$, then $W+\alpha$ is a right coset of W in V. Since V is an abelian group, therefore the right coset $W+\alpha$ will be equal to the left coset $\alpha+W$. Thus we can say that $W+\alpha$ is a coset of W in V. Let V/W denote the set of all cosets of W in V i e , let $V/W=\{W+\alpha : \alpha \in V\}$.

Also we know that if $W+\alpha$ and $W+\beta$ are two cosets of W in V, then $W+\alpha=W+\beta \Leftrightarrow \alpha-\beta \in W$.

Now we shall give a vector space structure to the set V/W over the same field F. For this we shall have to define addition in V/W *i.e.*, addition of cosets of W in V and multiplication of a coset by an element of F *i.e.*, scalar multiplication.

Theorem. *If W is any subspace of a vector space $V(F)$, then the set V/W of all cosets $W+\alpha$ where α is any arbitrary element of V, is a vector space over F for the addition and scalar multiplication compositions defined as follows :*

$$(W+\alpha)+(W+\beta)=W+(\alpha+\beta)$$

and $\quad a(W+\alpha)=W+a\alpha,\ a \in F.$ (Delhi 1970)

Proof. Since $W+(\alpha+\beta) \in V/W$ and also $W+a\alpha \in V/W$, therefore V/W is closed with respect to addition of cosets and scalar multiplication as defined above. Now first of all we shall show that these two compositions are well defined.

Let $\quad W+\alpha=W+\alpha',\ \alpha, \alpha' \in V$

and $\quad W+\beta=W+\beta',\ \beta, \beta' \in V$.

We have $W+\alpha=W+\alpha' \Rightarrow \alpha-\alpha' \in W$

and $\quad W+\beta=W+\beta' \Rightarrow \beta-\beta' \in W$.

Now W is a subspace, therefore

$$\alpha-\alpha' \in W,\ \beta-\beta' \in W$$

$\Rightarrow (\alpha-\alpha')+(\beta-\beta') \in W \Rightarrow (\alpha+\beta)-(\alpha'+\beta') \in W$

$\Rightarrow W+(\alpha+\beta)=W+(\alpha'+\beta')$

$\Rightarrow (W+\alpha)+(W+\beta)=(W+\alpha')+(W+\beta')$.

Therefore addition in V/W is well defined.

Again $a \in F,\ \alpha-\alpha' \in W \Rightarrow a(\alpha-\alpha') \in W$

$\Rightarrow a\alpha-a\alpha' \in W \Rightarrow W+a\alpha=W+a\alpha' \Rightarrow a(W+\alpha)=a(W+\alpha')$.

$\therefore$ scalar multiplication in V/W is also well defined.

Commutativity of addition. Let $W+\alpha$, $W+\beta$ be any two elements of V/W. Then

$(W+\alpha)+(W+\beta)=W+(\alpha+\beta)=W+(\beta+\alpha)=(W+\beta)+(W+\alpha)$.

Associativity of addition. Let $W+\alpha$, $W+\beta$, $W+\gamma$ be any three elements of V/W. Then

$(W+\alpha)+[(W+\beta)+(W+\gamma)]=(W+\alpha)+[W+(\beta+\gamma)]$

$=W+[\alpha+(\beta+\gamma)]=W+[(\alpha+\beta)+\gamma]$

$=[W+(\alpha+\beta)]+(W+\gamma)=[(W+\alpha)+(W+\beta)]+(W+\gamma)$.

Existence of additive identity. If $\mathbf{0}$ is the zero vector of V, then $W+\mathbf{0}=W \in V/W$. If $W+\alpha$ is any element of V/W, then

$(W+0)+(W+\alpha)=W+(0+\alpha)=W+\alpha$.

$\therefore$ $W+0=W$ is the additive identity.

Existence of additive inverse. *If $W+\alpha$ is any element of V/W,* then $W-\alpha \in V/W$. Also we have

$$(W+\alpha)+(W-\alpha)=W+(\alpha-\alpha)=W+0=W.$$

$\therefore$ $W-\alpha$ is the additive inverse of $W+\alpha$.

Thus V/W is an abelian group with respect to addition composition. Further we observe that if

$$a, b \in F \text{ and } W+\alpha, W+\beta \in V/W, \text{ then}$$

1. $a[(W+\alpha)+(W+\beta)]=a[W+(\alpha+\beta)]=W+a(\alpha+\beta)$
$=(W+a\alpha)+(W+a\beta)=a(W+\alpha)+a(W+\beta)$.
2. $(a+b)(W+\alpha)=W+(a+b)\alpha=W+(a\alpha+b\alpha)$
$=(W+a\alpha)+(W+b\alpha)=a(W+\alpha)+b(W+\alpha)$.
3. $(ab)(W+\alpha)=W+(ab)\alpha=W+a(b\alpha)$
$=a(W+b\alpha)=a[b(W+\alpha)]$.
4. $1(W+\alpha)=W+1\alpha=W+\alpha$.

$\therefore$ V/W is a vector space over F for these two compositions. The vector space V/W is called the **Quotient Space** of V relative to W. The coset W is the zero vector of this vector space.

Dimension of a Quotient Space.

Theorem. *If W be a subspace of a finite dimensional vector space $V(F)$, then dim V/W=dim V−dim W.*

(I.A.S. 1972 ; Aligarh 65 ; Delhi Hon's. 71)

Proof. Let m be the dimension of the subspace W of the vector space $V(F)$. Let $S=\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ be a basis of W. Since S is a linearly independent subset of V, therefore it can be extended to form a basis of V. Let $S'=\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_l\}$ be a basis of V. Then dim $V=m+l$.

$$\therefore \quad \dim V-\dim W=(m+l)-m=l.$$

So we should prove that dim $V/W=l$.

We claim that the set of l cosets

$$S_1=\{W+\beta_1, W+\beta_2, \ldots, W+\beta_l\}$$

is a basis of V/W.

First we show that S_1 is linearly independent. The zero vector of V/W is W.

Let $a_1(W+\beta_1)+a_2(W+\beta_2)+\ldots+a_l(W+\beta_l)=W$

$$\Rightarrow (W+a_1\beta_1)+(W+a_2\beta_2)+\ldots+(W+a_l\beta_l)=W$$

$$\Rightarrow W+(a_1\beta_1+a_2\beta_2+\ldots+a_l\beta_l)=W$$

$\Rightarrow a_1\beta_1+a_2\beta_2+\ldots+a_l\beta_l \in W$

$\Rightarrow a_1\beta_1+a_2\beta_2+\ldots+a_l\beta_l=b_1\alpha_1+b_2\alpha_2+\ldots+b_m\alpha_m$

[$\because$ any vector in W can be expressed as a linear combination of its basis]

$\Rightarrow a_1\beta_1+a_2\beta_2+\ldots+a_l\beta_l-b_1\alpha_1-b_2\alpha_2-\ldots-b_m\alpha_m=0$

$\Rightarrow a_1=0,\ a_2=0, \ldots, a_l=0$ since the vectors $\beta_1, \beta_2, \ldots, \beta_l, \alpha_1, \alpha_2, \ldots, \alpha_m$ are linearly independent.

$\therefore$ The set S_1 is linearly independent.

Now to show that $L(S_1)=V/W$. Let $W+\alpha$ be any element of V/W. Then $\alpha \in V$ can be expressed as

$$\alpha=c_1\alpha_1+c_2\alpha_2+\ldots+c_m\alpha_m+d_1\beta_1+d_2\beta_2+\ldots+d_l\beta_l$$
$$=\gamma+d_1\beta_1+d_2\beta_2+\ldots+d_l\beta_l \text{ where}$$
$$\gamma=c_1\alpha_1+c_2\alpha_2+\ldots+c_m\alpha_m \in W.$$

So $W+\alpha=W+(\gamma+d_1\beta_1+d_2\beta_2+\ldots+d_l\beta_l)$

$=(W+\gamma)+d_1\beta_1+d_2\beta_2+\ldots+d_l\beta_l$

$=W+(d_1\beta_1+d_2\beta_2+\ldots+d_l\beta_l)$

[$\because$ $\gamma \in W \Rightarrow W+\gamma=W$]

$=(W+d_1\beta_1)+(W+d_2\beta_2)+\ldots+(W+d_l\beta_l)$

$=d_1(W+\beta_1)+d_2(W+\beta_2)+\ldots+d_l(W+\beta_l)$.

Thus any element $W+\alpha$ of V/W can be expressed as a linear combination of S_1. Therefore $V/W=L(S_1)$.

$\therefore$ S_1 is a basis of V/W and consequently $\dim V/W=l$.

Hence the theorem.

§ 15. Direct sum of spaces.

Vector space as a direct sum of subspaces.

Definition. *Let $V(F)$ be a vector space and let $W_1, W_2, \ldots, W_m$ be subspaces of V. Then V is said to be the direct sum of $W_1, W_2, \ldots, W_m$ if every element $\alpha \in V$ can be written in one and only one way as $\alpha=\alpha_1+\alpha_2+\ldots+\alpha_m$ where*

$\alpha_1 \in W_1, \alpha_2 \in W_2, \ldots, \alpha_m \in W_m$. **(Garhwal 1976)**

If a vector space $V(F)$ is a direct sum of its two subspaces W_1 and W_2 then we should have not only $V=W_1+W_2$ but also that each vector of V can be uniquely expressed as sum of an element of W_1 and an element of W_2. Symbolically the direct sum is represented by the notation $V=W_1 \oplus W_2$.

Example. Let $V_2(F)$ be the vector space of all ordered pairs of F. Then $W_1=\{(a, 0) : a \in F\}$ and $W_2=\{(0, b) : b \in F\}$ are two subspaces of V_2 (F). Obviously any element $(x, y) \in V_2$ (F) can be uniquely expressed as sum of an element of W_1 and an element of W_2. This unique expression is $(x, y)=(x, 0)+(0, y)$. Thus $V_2(F)$ is the direct sum of W_1 and W_2. Also we observe that the only element common to both W_1 and W_2 is the zero vector $(0, 0)$.

Disjoint subspaces. Definition. *Two subspaces* W_1 *and* W_2 *of the vector space* V (F) *are said to be disjoint if their intersection is the zero subspace i.e., if* $W_1 \cap W_2=\{0\}$.

Theorem. *The necessary and sufficient conditions for a vector space* V (F) *to be a direct sum of its two subspaces* W_1 *and* W_2 *are that*

$$(i) \quad V=W_1+W_2$$

and
$$(ii) \quad W_1 \cap W_2=\{0\}$$

i.e., W_1 *and* W_2 *are disjoint.* **(Kanpur 1970 ; Garhwal 76)**

Proof. The conditions are necessary.

Let V be a direct sum of its two subspaces W_1 and W_2. Then each element of V is expressible as sum of an element of W_1 and an element of W_2. Therefore we have $V=W_1+W_2$.

Let, if possible, $0 \neq \alpha \in W_1 \cap W_2$. Then $\alpha \in W_1$, $\alpha \in W_2$. Also $\alpha \in V$ and we can write

$$\alpha=0+\alpha \text{ where } 0 \in W_1, \ \alpha \in W_2$$

and
$$\alpha=\alpha+0 \text{ where } \alpha \in W_1, \ 0 \in W_2.$$

Thus $\alpha \in V$ can be expressed in at least two different ways as sum of an element of W_1 and an element of W_2. This contradicts the fact that V is a direct sum of W_1 and W_2. Hence 0 is the only vector common to both W_1 and W_2 *i.e.*, $W_1 \cap W_2=\{0\}$. Thus the conditions are necessary.

The conditions are sufficient.

Let $V=W_1+W_2$ and $W_1 \cap W_2=\{0\}$. Then to show that V is direct sum of W_1 and W_2.

$V=W_1+W_2 \Rightarrow$ that each element of V can be expressed as sum of an element of W_1 and an element of W_2. Now to show that this expression is unique. Let, if possible,

$$\alpha=\alpha_1+\alpha_2, \ \alpha \in V, \ \alpha_1 \in W_1, \ \alpha_2 \in W_2,$$

and
$$\alpha=\beta_1+\beta_2, \ \beta_1 \in W_1, \ \beta_2 \in W_2.$$

Then to show that $\alpha_1=\beta_1$ and $\alpha_2=\beta_2$.

We have $\alpha_1+\alpha_2=\beta_1+\beta_2 \Rightarrow \alpha_1-\beta_1=\beta_2-\alpha_2$.

Since W_1 is a subspace, therefore

$$\alpha_1 \in W_1,\ \beta_1 \in W_1 \Rightarrow \alpha_1 - \beta_1 \in W_1.$$

Similarly $\beta_2 - \alpha_2 \in W_2$.

$\therefore\ \alpha_1 - \beta_1 = \beta_2 - \alpha_2 \in W_1 \cap W_2$.

But $\mathbf{0}$ is the only vector which belongs to $W_1 \cap W_2$. Therefore $\alpha_1 - \beta_1 = \mathbf{0} \Rightarrow \alpha_1 = \beta_1$. Also $\beta_2 - \alpha_2 = \mathbf{0} \Rightarrow \alpha_2 = \beta_2$.

Thus each vector $\alpha \in V$ is uniquely expressible as sum of an element of W_1 and an element of W_2. Hence $V = W_1 \oplus W_2$.

Dimension of a direct sum. Theorem. *If a finite dimensional vector space V (F) is a direct sum of its two subspaces W_1 and W_2, then*

$$\dim V = \dim W_1 + \dim W_2.$$

Proof. Let $\dim W_1 = m$ and $\dim W_2 = l$. Also let the sets of vectors $S_1 = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ and $S_2 = \{\beta_1, \beta_2, \ldots, \beta_l\}$ be some bases of W_1 and W_2 respectively.

We have $\dim W_1 + \dim W_2 = m + l$.

In order to prove that $\dim V = \dim W_1 + \dim W_2$, we should therefore prove that $\dim V = m + l$. We claim that the set

$$S = S_1 \cup S_2 = \{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_l\}$$

is a basis of V.

First we show that the set S is linearly independent. Let

$$a_1\alpha_1 + a_2\alpha_2 + \ldots + a_m\alpha_m + b_1\beta_1 + b_2\beta_2 + \ldots + b_l\beta_l = \mathbf{0}$$

$$\Rightarrow a_1\alpha_1 + a_2\alpha_2 + \ldots + a_m\alpha_m = -(b_1\beta_1 + b_2\beta_2 + \ldots + b_l\beta_l).$$

Now $a_1\alpha_1 + a_2\alpha_2 + \ldots + a_m\alpha_m \in W_1$

and $\quad -(b_1\beta_1 + b_2\beta_2 + \ldots + b_l\beta_l) \in W_2$. Therefore

$$a_1\alpha_1 + a_2\alpha_2 + \ldots + a_m\alpha_m \in W_1 \cap W_2$$

and $\quad -(b_1\beta_1 + b_2\beta_2 + \ldots + b_l\beta_l) \in W_1 \cap W_2$.

But V is the direct sum of W_1 and W_2. Therefore $\mathbf{0}$ is the only vector belonging to $W_1 \cap W_2$. Then we have

$$a_1\alpha_1 + a_2\alpha_2 + \ldots + a_m\alpha_m = \mathbf{0},\ b_1\beta_1 + b_2\beta_2 + \ldots + b_l\beta_l = \mathbf{0}.$$

Since both the sets $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ and $\{\beta_1, \beta_2, \ldots, \beta_l\}$ are linearly independent, therefore we have

$$a_1 = 0,\ a_2 = 0, \ldots, a_m = 0,\ b_1 = 0,\ b_2 = 0, \ldots, b_l = 0.$$

Therefore S is linearly independent.

Now we shall show that $L(S) = V$. Let α be any element of V. Then

$\alpha =$ an element of $W_1 +$ an element of W_2

$=$ a linear combination of $S_1 +$ a linear combination of S_2

$=$ a linear combination of S.

$\therefore\ L(S) = V$.

$\therefore\ S$ is a basis of V. Therefore $\dim V = m + l$.

Hence the theorem.

A sort of converse of this theorem is true. It has been proved in the following theorem :

Theorem. *Let V be a finite dimensional vector space and let W_1, W_2 be subspaces of V such that $V=W_1+W_2$ and*

$$\dim V=\dim W_1+\dim W_2. \text{ Then } V=W_1 \oplus W_2.$$

Proof. Let $\dim W_1=l$ and $\dim W_2=m$. Then $\dim V=l+m$.

Let $S_1=\{\alpha_1, \alpha_2, \ldots, \alpha_l\}$ be a basis of W_1 and $S_2=\{\beta_1, \beta_2, \ldots, \beta_m\}$ be a basis of W_2. We shall first show that $S_1 \cup S_2$ is a basis of V.

Let $\alpha \in V$. Since $V=W_1+W_2$, therefore we can write

$$\alpha=\gamma+\delta \text{ where } \gamma \in W_1, \delta \in W_2.$$

Now $\gamma \in W_1$ can be expressed as a linear combination of the elements of S_1 and $\delta \in W_2$ can be expressed as a linear combination of the elements of S_2. Therefore $\alpha \in V$ can be expressed as a linear combination of the elements of $S_1 \cup S_2$. Therefore

$$V=L\,(S_1 \cup S_2).$$

Since $\dim V=l+m$ and $L\,(S_1 \cup S_2)=V$, therefore the number of distinct elements in $S_1 \cup S_2$ cannot be less than $l+m$.

Thus $S_1 \cup S_2$ has $l+m$ distinct elements and therefore $S_1 \cup S_2$ is a basis of V. Therefore the set $\{\alpha_1, \alpha_2, \ldots, \alpha_l, \beta_1, \beta_2, \ldots, \beta_m\}$ is linearly independent.

Now we shall show that $W_1 \cap W_2=\{0\}$.

Let $\alpha \in W_1 \cap W_2$. Then $\alpha \in W_1$ $\alpha \in W_2$.

Therefore $\alpha=a_1\alpha_1+a_2\alpha_2+\ldots+a_l\alpha_l$ and $\alpha=b_1\beta_1+b_2\beta_2+\ldots+b_m\beta_m$ for some a's and b's$\in F$.

$$\therefore \quad a_1\alpha_1+a_2\alpha_2+\ldots+a_l\alpha_l=b_1\beta_1+b_2\beta_2+\ldots+b_m\beta_m$$
$$\Rightarrow a_1\alpha_1+a_2\alpha_2+\ldots+a_l\alpha_l-b_1\beta_1-b_2\beta_2-\ldots-b_m\beta_m=0$$
$$\Rightarrow a_1=0, a_2=0, \ldots, a_l=0, b_1=0, b_2=0, \ldots, b_m=0 \Rightarrow \alpha=0.$$
$$\therefore \quad W_1 \cap W_2=\{0\}.$$

Now $V=W_1+W_2$ and $W_1 \cap W_2=\{0\} \Rightarrow V=W_1 \oplus W_2$.

Complementary subspaces. Definition. *Let $V\,(F)$ be a vector space and W_1, W_2 be two subspaces of V. Then the subspace W_2 is called a complement of W_1 in V if V is the direct sum of W_1 and W_2.*

Existence of complementary subspaces. Theorem. *Corresponding to each subspace W_1 of a finite dimensional vector space $V\,(F)$, there exists a subspace W_2 such that V is the direct sum of W_1 and W_2.* **(I.A.S. 1972 ; Meerut 68)**

Proof. Let $\dim W_1=m$. Let the set $S_1=\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ be a basis of W_1. Since S_1 is a linearly independent subset of V, therefore S_1 can be extended to form a basis of V. Let the set

$S=\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_l\}$ be a basis of V. Let W_2 be the subspace of V generated by the set $S_2=\{\beta_1, \beta_2, \ldots, \beta_l\}$.

We shall prove that V is the direct sum of W_1 and W_2. For this we shall prove that $V=W_1+W_2$ and $W_1 \cap W_2=\{0\}$.

Let α be any element of V. Then we can express

$\alpha=$a linear combination of S

$=$a linear combination of S_1+a linear combination of S_2

$=$an element of W_1+an element of W_2.

$\therefore \quad V=W_1+W_2$.

Again let $\beta \in W_1 \cap W_2$. Then β can be expressed as a linear combination of S_1 and also a linear combination of S_2. So we have

$$\beta=a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m=b_1\beta_1+b_2\beta_2+\ldots+b_l\beta_l.$$

$$\therefore \quad a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m-b_1\beta_1-b_2\beta_2-\ldots-b_l\beta_l=0$$

$\Rightarrow a_1=0,\ a_2=0,\ \ldots,\ a_m=0,\ b_1=0,\ b_2=0,\ \ldots,\ b_l=0$ since $\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_l$ are linearly independent.

$\therefore \quad \beta=0$ (zero vector).

Thus $W_1 \cap W_2=\{0\}$. Hence V is the direct sum of W_1 and W_2.

§ **16. Coordinates.** Let $V(F)$ be a finite dimensional vector space. Let $B=\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be an ordered basis for V. By an ordered basis we mean that the vectors of B have been enumerated in some well-defined way *i.e.*, the vectors occupying the first, second, ..., n^{th} places in the set B are fixed.

Let $\alpha \in V$. Then there exists a unique n-tuple $(x_1, x_2, \ldots, x_n)$ of scalars such that $\alpha=x_1\alpha_1+x_2\alpha_2+\ldots+x_n\alpha_n$.

The n-tuple $(x_1, x_2, \ldots, x_n)$ is called the n-tuple of **coordinates of α relative to the ordered basis B.** The scalar x_i is called the i^{th} **coordinate of α relative to the ordered basis B.**

It should be noted that for the same basis set B, the coordinates of the vector α are unique only with respect to particular ordering of B. The basis set B can be ordered in several ways. The co-ordinates of α may change with change in ordering of B.

Example. *Show that the set* $S=\{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$ *is a basis of* $\mathbf{R}^3(\mathbf{R})$ *where* $\mathbf{R}$ *is the field of real numbers. Hence find the coordinates of the vector* (a, b, c) *with respect to the above basis.*

Solution. The dimension of the vector space $\mathbf{R}^3(\mathbf{R})$ is 3. If the set S is linearly independent, then S will form a basis of $\mathbf{R}^3(\mathbf{R})$ because S contains 3 vectors. Let x, y, z be scalars in $\mathbf{R}$ such that

$$x(1, 0, 0)+y(1, 1, 0)+z(1, 1, 1)=\mathbf{0}=(0, 0, 0)$$

$$\Rightarrow (x+y+z, y+z, z)=(0, 0, 0)$$

$\Rightarrow x+y+z=0,\ y+z=0,\ z=0 \Rightarrow x=0,\ y=0,\ z=0$

$\Rightarrow$ the set S is linearly independent.

$\therefore$ S is a basis of $\mathbf{R}^3(\mathbf{R})$.

Now to find the coordinates of (a, b, c) with respect to the ordered basis S. Let p, q, r be scalars in $\mathbf{R}$ such that

$$(a, b, c)=p(1, 0, 0)+q(1, 1, 0)+r(1, 1, 1)$$

$$\Rightarrow (a, b, c)=(p+q+r,\ q+r,\ r)$$

$$\Rightarrow p+q+r=a,\ q+r=b,\ r=c \Rightarrow r=c,\ q=b-c,\ p=a-b.$$

Hence the coordinates of the vector (a, b, c) are (p, q, r) *i.e.*, $(a-b, b-c, c)$.

Exercises

1. Construct three subspaces W_1, W_2, W_3 of a vector space V so that $V=W_1 \oplus W_2=W_1 \oplus W_3$ but $W_2 \neq W_3$.

   Hint : Take $V=\mathbf{R}^2$, $W_1=\{(a, 0) : a \in \mathbf{R}\}$, $W_2=\{(0, a) : a \in \mathbf{R}\}$, and $W_3=\{(a, a) : a \in \mathbf{R}\}$.

2. Find the co-ordinates of the vector $(2, 1, -6)$ of $\mathbf{R}^3$ relative to the basis $\alpha_1=(1, 1, 2)$, $\alpha_2=(3, -1, 0)$, $\alpha_3=(2, 0, -1)$.

   Ans. $(-\frac{7}{8}, -\frac{15}{8}, \frac{17}{4})$.

7

Vector Spaces (Continued)

§ 1. **Linear Transformations.** In the preceding chapter we have defined **linear transformation** or **vector space homomorphism.** For the sake of convenience we repeat the definition here again.

Linear transformation. Definition. *Let U and V be two vector spaces over the same field F. A linear transformation from U into V is a function T from U into V such that*

$$T(a\alpha+b\beta)=aT(\alpha)+bT(\beta)$$

for all α, β *in U and for all a, b in F.*

Linear Operator. Definition. *Let V be a vector space over the field F. A linear operator on V is a function T from V into V such that*

$$T(a\alpha+b\beta)=aT(\alpha)+bT(\beta)$$

for all α, β *in V and for all a, b in F.*

Thus T is a linear operator on V if T is a linear transformation from V into V itself.

Linear functional. Definition. *Let V be a vector space over the field F. A function f from V into F is said to be a linear functional on V if*

$$f(a\alpha+b\beta)=af(\alpha)+bf(\beta)$$

for all $\alpha, \beta \in V$ *and for all* $a, b \in F$.

We know that if F is any field, then F can be regarded as a vector space over F Therefore f is a linear functional on V if f is a linear transformation from V into F.

§ 2. Linear transformations as vectors.

Let $L(U, V)$ be the set of all linear transformations of a vector space $U(F)$ into a vector space $V(F)$. Sometimes we denote this set by Hom (U, V). Now we want to impose a vector space structure on the set $L(U, V)$ over the same field F For this purpose we shall have to suitably define addition in $L(U, V)$ and scalar multiplication in $L(U, V)$ over F.

Theorem 1. *Let U and V be vector spaces over the field F. Let T_1 and T_2 be linear transformations from U into V. The function T_1+T_2 defined by*

$$(T_1+T_2)(\alpha)=T_1(\alpha)+T_2(\alpha)\ \forall\, \alpha\in U$$

is a linear transformation from U into V. If c is any element of F, the function (cT) defined by

$$(cT)(\alpha)=cT(\alpha)\ \forall\, \alpha\in U$$

is a linear transformation from U into V. The set $L(U, V)$ of all linear transformations from U into V, together with the addition and scalar multiplication defined above, is a vector space over the field F.

(Kanpur 1969)

Proof. Suppose T_1 and T_2 are linear transformations from U into V and we define T_1+T_2 as follows :

$$(T_1+T_2)(\alpha)=T_1(\alpha)+T_2(\alpha)\ \forall\, \alpha\in U. \qquad \ldots(1)$$

Since $T_1(\alpha)+T_2(\alpha)\in V$, therefore T_1+T_2 is a function from U into V.

Let $a, b\in F$ and $\alpha, \beta\in U$. Then

$(T_1+T_2)(a\alpha+b\beta)=T_1(a\alpha+b\beta)+T_2(a\alpha+b\beta)$ [by (1)]

$=[aT_1(\alpha)+bT_1(\beta)]+[aT_2(\alpha)+bT_2(\beta)]$

[$\because$ T_1 and T_2 are linear transformations]

$=a[T_1(\alpha)+T_2(\alpha)]+b[T_1(\beta)+T_2(\beta)]$ [$\because$ V is a vector space]

$=a(T_1+T_2)(\alpha)+b(T_1+T_2)(\beta)$ [by (1)]

$\therefore$ T_1+T_2 is a linear transformation from U into V. Thus $T_1, T_2\in L(U, V) \Rightarrow T_1+T_2\in L(U, V)$. Therefore $L(U, V)$ is closed with respect to addition defined on it.

Again let $T\in L(U, V)$ and $c\in F$. Let us define cT as follows :

$$(cT)(\alpha)=cT(\alpha)\ \forall\, \alpha\in U. \qquad \ldots(2)$$

Since $cT(\alpha)\in V$, therefore cT is a function from U into V.

Let $a, b\in F$ and $\alpha, \beta\in U$. Then

$(cT)(a\alpha+b\beta)=cT(a\alpha+b\beta)$ [by (2)]

$=c[aT(\alpha)+bT(\beta)]$ [$\because$ T is a linear transformation]

$=c[aT(\alpha)]+c[bT(\beta)]=(ca)\,T(\alpha)+(cb)\,T(\beta)$

$=(ac)\,T(\alpha)+(bc)\,T(\beta)=a[cT(\alpha)]+b[cT(\beta)]$

$=a[(cT)(\alpha)]+b[(cT)(\beta)]$.

$\therefore$ cT is a linear transformation from U into V. Thus $T\in L(U, V)$ and $c\in F \Rightarrow cT\in L(U, V)$. Therefore $L(U, V)$ is closed with respect to scalar multiplication defined in it.

Associativity of addition in $L(U, V)$.

Let $T_1, T_2, T_3 \in L(U, V)$. If $\alpha \in U$, then

$[T_1+(T_2+T_3)](\alpha)=T_1(\alpha)+(T_2+T_3)(\alpha)$

[by (1) *i.e.*, by def. of addition in $L(U, V)$]

$=T_1(\alpha)+[T_2(\alpha)+T_3(\alpha)]$ [by (1)]

$=[T_1(\alpha)+T_2(\alpha)]+T_3(\alpha)$ [$\because$ addition in V is associative]

$=(T_1+T_2)(\alpha)+T_3(\alpha)$ [by (1)]

$=[(T_1+T_2)+T_3](\alpha)$ [by (1)]

$\therefore \quad T_1+(T_2+T_3)=(T_1+T_2)+T_3$

[by def. of equality of two functions]

Commutativity of addition in $L(U, V)$. Let $T_1, T_2 \in L(U, V)$. If α is any element of U, then

$(T_1+T_2)(\alpha)=T_1(\alpha)+T_2(\alpha)$ [by (1)]

$=T_2(\alpha)+T_1(\alpha)$ [$\because$ addition in V is commutative]

$=(T_2+T_1)(\alpha)$ [by (1)]

$\therefore \quad T_1+T_2=T_2+T_1$ [by def. of equality of two functions]

Existence of additive identity in $L(U, V)$. Let $\hat{0}$ be the zero function from U into V *i.e.*, $\hat{0}(\alpha)=0 \in V \ \forall \ \alpha \in U$.

It can be easily seen that $\hat{0}$ is a linear transformation from U into V. Thus $\hat{0} \in L(U, V)$. If $T \in L(U, V)$ and $\alpha \in U$, we have

$(\hat{0}+T)(\alpha)=\hat{0}(\alpha)+T(\alpha)$ [by (1)]

$=0+T(\alpha)$ [by def. of $\hat{0}$]

$=T(\alpha)$ [0 being additive identity in V]

$\therefore \quad \hat{0}+T=T \ \forall \ T \in L(U, V)$.

$\therefore \quad \hat{0}$ is the additive identity in $L(U, V)$.

Existence of additive inverse of each element in $L(U, V)$.

Let $T \in L(U, V)$. Let us define $-T$ as follows :

$$(-T)(\alpha)=-T(\alpha) \ \forall \ \alpha \in U.$$

It can be easily seen that $-T$ is a linear transformation from U into V. Thus $-T \in L(U, V)$. If $\alpha \in U$, we have

$(-T+T)(\alpha)=(-T)(\alpha)+T(\alpha)$ [by def. of addition in $L(U, V)$]

$=-T(\alpha)+T(\alpha)$ [by def. of $-T$]

$=0 \in V$

$=\hat{0}(\alpha)$ [by def. of $\hat{0}$]

$\therefore \quad -T+T=\hat{0}$ for every $T \in L(U, V)$.

Thus each element in $L(U, V)$ possesses additive inverse.

Therefore $L(U, V)$ is an **abelian group** with respect to addition defined in it.

Further we make the following observations :

(i) Let $c \in F$ and $T_1, T_2 \in L(U, V)$. If α is any element in U, we have

$[c(T_1+T_2)](\alpha)=c[(T_1+T_2)(\alpha)]$

[by (2) *i.e.*, by def. of scalar multiplication in $L(U, V)$]

$=c[T_1(\alpha)+T_2(\alpha)]$ [by (1)]

$=cT_1(\alpha)+cT_2(\alpha)$

[$\because$ $c \in F$ and $T_1(\alpha), T_2(\alpha) \in V$ which is a vector space]

$=(cT_1)(\alpha)+(cT_2)(\alpha)$ [by (2)]

$=(cT_1+cT_2)(\alpha)$ [by (1)]

$\therefore \quad c(T_1+T_2)=cT_1+cT_2$.

(ii) Let $a, b \in F$ and $T \in L(U, V)$. If $\alpha \in U$, we have

$[(a+b)T](\alpha)=(a+b)T(\alpha)$ [by (2)]

$=aT(\alpha)+bT(\alpha)$ [$\because$ V is a vector space]

$=(aT)(\alpha)+(bT)(\alpha)$ [by (2)]

$=(aT+bT)(\alpha)$ [by (1)]

$\therefore \quad (a+b)T=aT+bT$.

(iii) Let $a, b \in F$ and $T \in L(U, V)$. If $\alpha \in U$, we have

$[(ab)T](\alpha)=(ab)T(\alpha)$ [by (2)]

$=a[bT(\alpha)]$ [$\because$ V is a vector space]

$=a[(bT)(\alpha)]$ [by (2)]

$=[a(bT)](\alpha)$ [by (2)]

$\therefore \quad (ab)T=a(bT)$.

(iv) Let $1 \in F$ and $T \in L(U, V)$. If $\alpha \in U$, we have

$(1T)(\alpha)=1\,T(\alpha)$ [by (2)]

$=T(\alpha)$ [$\because$ V is a vector space]

$\therefore \quad 1T=T$.

Hence $L(U, V)$ is a vector space over the field F.

Note. If in place of the vector space V, we take U, then we observe that the set of all linear operators on U forms a vector space with respect to addition and scalar multiplication defined above.

Similarly if in place of the vector space V we take the field F as a vector space, then we observe that the set of all linear functionals on U forms a vector space with respect to addition and scalar multiplication defined above.

Dual space. Definition. *Let V be a vector space over the field F. Then the set V' of all linear functionals on V is also a vector space over the field F. The vector space V' is called the dual space of V.*

(Meerut 1971)

Sometimes V^* and $\hat{V}$ are also used to denote the dual space of V.

Dimension of $L(U, V)$. Now we shall prove that if $U(F)$ and $V(F)$ are finite dimensional, then the vector space of all linear transformations from U into V is also finite dimensional. For this purpose we shall require an **important result** which we prove in the following theorem :

Theorem 2. *Let U be a finite dimensional vector space over the field F and let $B=\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be an ordered basis for U. By an ordered basis we mean that the vectors of B have been enumerated in some well defined way, i.e., the vectors occupying the first, second,..., n^{th} places in the set B are fixed. Let V be a vector space over the same field F and let $\beta_1, \ldots, \beta_n$ be any n vectors in V. Then there exists a unique linear transformation T from U into V such that*

$$T(\alpha_i)=\beta_i,\ i=1, 2, \ldots, n. \qquad \textbf{(Sambalpur 1977)}$$

Proof. Existence of T. Let $\alpha \in U$. Since $B=\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a basis for U, therefore there exist unique scalars $x_1, x_2, \ldots, x_n$ such that $\alpha=x_1\alpha_1+x_2\alpha_2+\ldots+x_n\alpha_n$.

For this vector α, let us define

$$T(\alpha)=x_1\beta_1+x_2\beta_2+\ldots+x_n\beta_n.$$

Obviously $T(\alpha)$ as defined above is a unique element of V. Therefore T is a well-defined rule for associating with each vector α in U a unique vector $T(\alpha)$ in V. Thus T is a function from U into V.

The unique representation of $\alpha_i \in U$ as a linear combination of the vectors belonging to the basis B is

$$\alpha_i=0\alpha_1+0\alpha_2+\ldots+1\alpha_i+0\alpha_{i+1}+\ldots+0\alpha_n.$$

Therefore according to our definition of T, we have

$$T(\alpha_i)=0\beta_1+0\beta_2+\ldots+1\beta_i+0\beta_{i+1}+\ldots+0\beta_n$$

i.e., $\quad T(\alpha_i)=\beta_i,\ i=1, 2, \ldots, n.$

Now to show that T is a linear transformation.

Let $a, b \in F$ and $\alpha, \beta \in U$. Let

$\alpha=x_1\alpha_1+\ldots+x_n\alpha_n$ and $\beta=y_1\alpha_1+\ldots+y_n\alpha_n$.

Then $T(a\alpha+b\beta)=T[a(x_1\alpha_1+\ldots+x_n\alpha_n)+b(y_1\alpha_1+\ldots+y_n\alpha_n)]$

$=T[(ax_1+by_1)\alpha_1+\ldots+(ax_n+by_n)\alpha_n]$

$=(ax_1+by_1)\beta_1+\ldots+(ax_n+by_n)\beta_n$ [by def. of T]

$=a(x_1\beta_1+\ldots+x_n\beta_n)+b(y_1\beta_1+\ldots+y_n\beta_n)$

$=a\,T(\alpha)+b\,T(\beta).$ [by def. of T]

$\therefore$ T is a linear transformation from U into V. Thus there exists a linear transformation from U into V such that

$$T(\alpha_i)=\beta_i,\ i=1,2,\ldots,n.$$

Uniqueness of T. Let T' be a linear transformation from U into V such that $T'(\alpha_i)=\beta_i$, $i=1,2,\ldots,n$. For the vector

$\alpha=x_1\alpha_1+\ldots+x_n\alpha_n\in U$, we have

$T'(\alpha)=T'(x_1\alpha_1+\ldots+x_n\alpha_n)$

$=x_1T'(\alpha_1)+\ldots+x_nT'(\alpha_n)$ [$\because$ T' is a linear transformation]

$=x_1\beta_1+\ldots+x_n\beta_n$ [by def of T']

$=T(\alpha).$ [by def. of T]

Thus $T'(\alpha)=T(\alpha)\ \forall\ \alpha\in U$. Therefore $T'=T$.

This shows the uniqueness of T.

Note. From this theorem we conclude that if T is a linear transformation from a finite dimensional vector space $U(F)$ into a vector space $V(F)$, then T is completely defined if we mention under T the images of the elements of a basis set of U. If S and T are two linear transformations from U into V such that $S(\alpha_i)=T(\alpha_i)\ \forall\ \alpha_i$ belonging to a basis of U, then

$$S(\alpha)=T(\alpha)\ \forall\ \alpha\in U,\ i.e.,\ S=T.$$

Thus two linear transformations from U into V are equal if they agree on a basis of U.

Corollary. *Let V be an n-dimensional vector space over the field F and let* $B=\{\alpha_1,\ldots,\alpha_n\}$ *be an ordered basis for V. If* $\{x_1,\ldots,x_n\}$ *is any ordered set of n scalars, then there exists a unique linear functional f on V, such that*

$$f(\alpha_i)=x_i,\ i=1,2,\ldots,n.$$

Theorem 3. *Let U be an n-dimensional vector space over the field F, and let V be an m-dimensional vector space over F. Then the vector space L (U, V) of linear transformations from U into V is also finite dimensional and is of dimension mn.* **(Meerut 1971)**

Proof. Let $B=\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ and $B'=\{\beta_1,\beta_2,\ldots,\beta_m\}$ be ordered bases for U and V respectively. By theorem 2, there exists a unique linear transformation T_{11} from U into V such that

$T_{11}(\alpha_1)=\beta_1,\ T_{11}(\alpha_2)=0,\ldots,\ T_{11}(\alpha_n)=0$ where

$\beta_1,0,\ldots,0$ are vectors in V.

In fact, for each pair of integers (p, q) with $1 \leqslant p \leqslant m$ and $1 \leqslant q \leqslant n$, there exists a unique linear transformation T_{pq} from U into V such that

$$T_{pq}(\alpha_i) = \begin{cases} 0, & \text{if } i \neq q \\ \beta_p, & \text{if } i = q \end{cases}$$

i.e., $$T_{pq}(\alpha_i) = \delta_{iq}\, \beta_p, \qquad \ldots(1)$$

where $\delta_{iq} \in F$ is Kronecker delta *i.e.*, $\delta_{iq} = 1$ if $i = q$ and $\delta_{iq} = 0$ if $i \neq q$.

Since p can be any of $1, 2, \ldots, m$ and q any of $1, 2, \ldots, n$, there are mn such T_{pq}'s. Let B_1 denote the set of these mn transformations T_{pq}'s. We shall show that B_1 is a basis for $L(U, V)$.

(i) First we shall show that $L(U, V)$ is a linear span of B_1.

Let $T \in L(U, V)$. Since $T(\alpha_1) \in V$ and any element in V is a linear combination of $\beta_1, \beta_2, \ldots, \beta_m$, therefore

$$T(\alpha_1) = a_{11}\beta_1 + a_{21}\beta_2 + \ldots + a_{m1}\beta_m$$

for some $a_{11}, a_{21}, \ldots, a_{m1} \in F$. In fact for each i, $1 \leqslant i \leqslant n$,

$$T(\alpha_i) = a_{1i}\beta_1 + a_{2i}\beta_2 + \ldots + a_{mi}\beta_m = \sum_{p=1}^{m} a_{pi}\beta_p. \qquad \ldots(2)$$

Now consider $S = \sum_{p=1}^{m} \sum_{q=1}^{n} a_{pq} T_{pq}$.

Obviously S is a linear combination of elements of B_1 which is a subset of $L(U, V)$. Since $L(U, V)$ is a vector space, therefore $S \in L(U, V)$ *i.e.*, S is also a linear transformation from U into V. We shall show that $S = T$.

Let us compute $S(\alpha_i)$ where α_i is any vector in the basis B of U. We have

$$S(\alpha_i) = \left[\sum_{p=1}^{m} \sum_{q=1}^{n} a_{pq} T_{pq}\right](\alpha_i) = \sum_{p=1}^{m} \sum_{q=1}^{n} a_{pq} T_{pq}(\alpha_i)$$

$$\sum_{p=1}^{m} \sum_{q=1}^{n} a_{pq}\, \delta_{iq}\, \beta_p \qquad \text{[From (1)]}$$

$= \sum_{p=1}^{m} a_{pi}\, \beta_p$ [On summing with respect to q. Remember that $\delta_{iq} = 1$ when $q = i$ and $\delta_{iq} = 0$ when $q \neq i$]

$= T(\alpha_i)$. [From (2)]

Thus $S(\alpha_i) = T(\alpha_i)\ \forall\ \alpha_i \in B$. Therefore S and T agree on a basis of U. So we must have $S=T$. Thus T is also a linear combination of the elements of B_1. Therefore $L(U, V)$ is a linear span of B_1.

(ii) Now we shall show that B_1 is linearly independent. Let for b_{pq}'s $\in F$,

$$\sum_{p=1}^{m} \sum_{q=1}^{n} b_{pq} T_{pq} = \hat{0} \text{ i.e., zero vector of } L(U, V)$$

$$\Rightarrow \left[\sum_{p=1}^{m} \sum_{q=1}^{n} b_{pq} T_{pq}\right](\alpha_i) = \hat{0}(\alpha_i) \ \forall \ \alpha_i \in B$$

$$\Rightarrow \sum_{p=1}^{m} \sum_{q=1}^{n} b_{pq} T_{pq}(\alpha_i) = 0 \in V \ [\because \hat{0} \text{ is zero transformation}]$$

$$\Rightarrow \sum_{p=1}^{m} \sum_{q=1}^{n} b_{pq} \delta_{iq} \beta_p = 0 \Rightarrow \sum_{p=1}^{m} b_{pi} \beta_p = 0$$

$$\Rightarrow b_{1i}\beta_1 + b_{2i}\beta_2 + \dots + b_{mi}\beta_m = 0, \ 1 \leqslant i \leqslant n$$

$$\Rightarrow b_{1i} = 0, b_{2i} = 0, \dots, b_{mi} = 0, \ 1 \leqslant i \leqslant n$$

$[\because \beta_1, \beta_2, \dots, \beta_m$ are linearly independent$]$

$\Rightarrow b_{pq}=0$ where $1 \leqslant p \leqslant m$ and $1 \leqslant q \leqslant n$

$\Rightarrow B_1$ is linearly independent.

Therefore B_1 is a basis of $L(U, V)$.

$\therefore$ dim $L(U, V)=$ the number of elements in $B_1=mn$.

Corollary 1. *The vector space $L(U, U)$ of all linear operators on an n-dimensional vector space U is of dimension n^2.*

Corollary 2. *The vector space V' of all linear functionals on an n-dimensional vector space $V(F)$ is of dimension n.*

Proof. Since $F(F)$ is a vector space of dimension 1, therefore the result is obvious.

Note. Suppose $U(F)$ is an n-dimensional vector space and $V(F)$ is an m-dimensional vector space. If $U \neq \{0\}$ and $V \neq \{0\}$, then $n \geqslant 1$ and $m \geqslant 1$. Therefore $L(U, V)$ does not just consist of the element $\hat{0}$, because dimension of $L(U, V)$ is $mn \geqslant 1$.

§ 3. Dual basis.

Theorem 1. *Let V be an n-dimensional vector space over the field F, and let $B=\{\alpha_1, \dots, \alpha_n\}$ be a basis for V. Then there is a uniquely determined basis $B'=\{f_1, \dots, f_n\}$ for V' such that $f_i(\alpha_j)=\delta_{ij}$. Consequently the dual space of an n-dimensional space is n dimensional.*

The basis B' is called the **dual basis** *of B.*

Proof. $B=\{\alpha_1,\ldots,\alpha_n\}$ is an ordered basis for V. Therefore by a corollary to theorem 2 on page 454 there exists a unique linear functional f_1 on V such that

$$f_1(\alpha_1)=1, f_1(\alpha_2)=0,\ldots, f_1(\alpha_n)=0$$

where $\{1, 0,\ldots, 0\}$ is an ordered set of n scalars.

In fact, for each $i=1, 2,\ldots, n$, there exists a unique linear functional f_i on V such that

$$f_i(\alpha_j)=\begin{cases}0 & \text{if } i\neq j\\ 1 & \text{if } i=j\end{cases}$$

i.e., $$f_i(\alpha_j)=\delta_{ij}, \qquad \ldots(1)$$

where $\delta_{ij}\in F$ is Kronecker delta *i.e.*, $\delta_{ij}=1$ if $i=j$ and $\delta_{ij}=0$ if $i\neq j$.

Let $B'=\{f_1,\ldots,f_n\}$. Then B' is a subset of V' containing n distinct elements of V'. We shall show that B' is a basis for V'.

First we shall show that B' is linearly independent.

Let $$c_1 f_1+c_2 f_2+\ldots+c_n f_n=\hat{0}$$

$$\Rightarrow (c_1 f_1+\ldots+c_n f_n)(\alpha)=\hat{0}(\alpha) \;\forall\; \alpha\in V$$

$$\Rightarrow c_1 f_1(\alpha)+\ldots+c_n f_n(\alpha)=0 \;\forall\; \alpha\in V \;[\because\; \hat{0}(\alpha)=0]$$

$$\Rightarrow \sum_{i=1}^{n} c_i f_i(\alpha)=0 \;\forall\; \alpha\in V$$

$$\Rightarrow \sum_{i=1}^{n} c_i f_i(\alpha_j)=0,\; j=1, 2,\ldots, n$$

[Putting $\alpha=\alpha_j$ where $j=1, 2,\ldots, n$]

$$\Rightarrow \sum_{i=1}^{n} c_i \delta_{ij}=0,\; j=1, 2,\ldots, n$$

$$\Rightarrow c_j=0,\; j=1, 2,\ldots, n$$

$\Rightarrow f_1, f_2,\ldots, f_n$ are linearly independent.

In the second place, we shall show that the linear span of B' is equal to V'.

Let f be any element of V'. The linear functional f will be completely determined if we define it on a basis for V_1. So let

$$f(\alpha_i)=a_i,\; i=1, 2,\ldots, n.$$

We shall show that $f=a_1 f_1+\ldots+a_n f_n$.

Let $\alpha=x_1\alpha_1+\ldots+x_n\alpha_n$ be any element in V. We have

$$f(\alpha)=f(x_1\alpha_1+\ldots+x_n\alpha_n)$$

$$=x_1 f(\alpha_1)+\dots+x_n f(\alpha_n) \qquad [\because f \text{ is linear}]$$

$$=x_1 a_1+\dots+x_n a_n. \qquad \dots(2)$$

Also $(a_1 f_1+\dots+a_n f_n)(\alpha)=a_1 f_1(\alpha)+\dots+a_n f_n(\alpha)$

$$= \sum_{i=1}^{n} a_i f_i(\alpha)$$

$$= \sum_{i=1}^{n} a_i f_i \left(\sum_{j=1}^{n} x_j \alpha_j \right) \qquad [\because \alpha=x_1\alpha_1+\dots+x_n\alpha_n]$$

$$= \sum_{i=1}^{n} a_i \sum_{j=1}^{n} x_j f_i(\alpha_j) \qquad [\because \text{ each } f_i \text{ is a linear functional}]$$

$$= \sum_{i=1}^{n} a_i \sum_{j=1}^{n} x_j \delta_{ij} \qquad [\text{From (1)}]$$

$$= \sum_{i=1}^{n} a_i x_i$$ [On summing with respect to j and remembering that $\delta_{ij}=1$ when $j=i$ and $\delta_{ij}=0$ when $j \neq i$]

$=a_1 x_1+\dots+a_n x_n=f(\alpha)$. [From (2)]

Thus we have $f(\alpha)=(a_1 f_1+\dots+a_n f_n)(\alpha) \; \forall \; \alpha \in V$.

$\therefore$ $f=a_1 f_1+\dots+a_n f_n$ *i.e.*, every element f in V' can be expressed as a linear combination of $f_1, \dots, f_n$.

$\therefore$ $V'=$linear span of B'. Hence B' is a basis for V'.

Now dim $V'=$number of distinct elements in $B'=n$.

Corollary. *If V is an n-dimensional vector space over the field F, then V is isomorphic to its dual space V′.*

Proof. We have dim $V'=$dim $V=n$.

$\therefore$ V is isomorphic to V'.

Theorem 2. *Let V be an n-dimensional vector space over the field F. If* α *is a non-zero vector in V, there exists a linear functional f on V such that* $f(\alpha) \neq 0$.

Proof. Since $\alpha \neq 0$, therefore $\{\alpha\}$ is a linearly independent subset of V. So it can be extended to form a basis for V. Thus there exists a basis $B=\{\alpha_1, \dots, \alpha_n\}$ for V such that $\alpha_1=\alpha$.

If $B'=\{f_1, \dots, f_n\}$ is the dual basis of V, then

$$f_1(\alpha)=f_1(\alpha_1)=1 \neq 0.$$

Thus there exists linear functional f_1 such that $f_1(\alpha) \neq 0$.

Corollary. *Let V be an n-dimensional vector space over the field F. If* $f(\alpha)=0 \; \forall \; f \in V'$, *then* $\alpha=0$.

Proof. Suppose $\alpha \neq 0$. Then there is a linear functional f on V such that $f(\alpha) \neq 0$. This contradicts the hypothesis that

$f(\alpha)=0 \ \forall \ f \in V'$. Hence we must have $\alpha=0$.

Theorem 3. *Let V be an n-dimensional vector space over the field F. If α, β are any two different vectors in V, then there exists a linear functional f on V such that $f(\alpha)\neq f(\beta)$.*

Proof. We have $\alpha\neq\beta \Rightarrow \alpha-\beta\neq 0$.

Now $\alpha-\beta$ is a non-zero vector in V. Therefore by theorm 2, there exists a linear functional f on V such that

$f(\alpha-\beta)\neq 0 \Rightarrow f(\alpha)-f(\beta)\neq 0 \Rightarrow f(\alpha)\neq f(\beta)$. Hence the result.

§ 4. Reflexivity.

Second dual space. We know that every vector space V possesses a dual space V' consisting of all linear functionals on V.

Now V' is also a vector space. Therefore it will also possess a dual space $(V')'$ consisting of all linear functionals on V'. This dual space of V' is called the **second dual space** of V and for the sake of simplicity we shall denote it by V''

If V is finite dimensional, then $\dim V=\dim V'=\dim V''$, showing that they are isomorphic to each other.

Now our aim is to establish an isomorphism of V onto V''.

Theorem 1. *Let V be a finite dimensional vector space over the field F. If α is any vector in V, the function L_α on V' defined by*

$$L_\alpha(f)=f(\alpha) \ \forall \ f\in V'$$

is a linear functional on V' i.e., $L_\alpha \in V''$.

Also the mapping $\alpha\rightarrow L_\alpha$ is an isomorphism of V onto V''.

(Kanpur 1969)

Proof. If $\alpha\in V$ and $f\in V'$, then $f(\alpha)$ is a unique element of F. Therefore the correspondence L_α defined by

$$L_\alpha(f)=f(\alpha) \ \forall \ f\in V' \quad \ldots(1)$$

is a function from V' into F.

Let $a, b\in F$ and $f, g\in V'$. Then

$L_\alpha(af+bg)=(af+bg)(\alpha)$ [From (1)]

$=(af)(\alpha)+(bg)(\alpha)$

$=af(\alpha)+bg(\alpha)$ [by scalar multiplication of linear functionals]

$=a[L_\alpha(f)]+b[L_\alpha(g)]$ [From (1)]

Therefore L_α is a linear functional on V' and thus $L_\alpha\in V''$.

Now let ψ be the function from V into V'' defined by

$$\psi(\alpha)=L_\alpha \ \forall \ \alpha\in V.$$

ψ is one-one. If $\alpha, \beta\in V$, then $\psi(\alpha)=\psi(\beta)$

$\Rightarrow L_\alpha=L_\beta \Rightarrow L_\alpha(f)=L_\beta(f) \ \forall \ f\in V'$

$\Rightarrow f(\alpha)=f(\beta) \ \forall \ f\in V'$ [from (1)]

$\Rightarrow f(\alpha)-f(\beta)=0 \ \forall f\in V' \Rightarrow f(\alpha-\beta)=0 \ \forall f\in V'$

$\Rightarrow \alpha-\beta=0$ [$\because$ by theorem 2 of § 3, if $\alpha-\beta\neq 0$, then $\exists$ a linear functional f on V such that $f(\alpha-\beta)\neq 0$. Here we have $f(\alpha-\beta)=0 \ \forall f\in V'$ and so $\alpha-\beta$ must be 0]

$\Rightarrow \alpha=\beta$. $\therefore$ ψ is one-one.

ψ is a linear transformation. Let $a, b \in F$ and $\alpha, \beta\in V$. Then

$\psi\ (a\alpha+b\beta)=L_{a\alpha+b\beta}$. [by def. of ψ]

For every $f\in V'$, we have

$$L_{a\alpha+b\beta}(f)=f(a\alpha+b\beta) \quad \text{[From (1)]}$$
$$=af(\alpha)+bf(\beta)$$
$$=aL_\alpha(f)+bL_\beta(f) \quad \text{[From (1)]}$$
$$=(aL_\alpha)\ (f)+(bL_\beta)\ (f)=(aL_\alpha+bL_\beta)\ (f).$$

$\therefore$ $L_{a\alpha+b\beta}=aL_\alpha+bL_\beta=a\psi(\alpha)+b\psi(\beta)$.

Thus $\psi(a\alpha+b\beta)=a\psi(\alpha)+b\psi(\beta)$.

$\therefore$ ψ is a linear transformation from V into V''.

Now ψ is a linear transformation from V into V''. We have dim $V=$dim V''. Therefore ψ is one-one implies that ψ must also be onto. Hence ψ is an isomorphism of V onto V''.

Note. The correspondence $\alpha\rightarrow L_\alpha$ as defined in the above theorem is called the **natural correspondence** between V and V''. It is important to note that the above theorem shows not only that V and V'' are isomorphic—this much is obvious from the fact that they have the same dimension—but that the natural correspondence is an isomorphism. This property of vector spaces is called **reflexivity**. Thus in the above theorem we have proved that *every finite dimensional vector space is reflexive.*

In future we shall identify V'' with V through the natural isomorphism $\alpha\leftrightarrow L_\alpha$. We shall say that the element L of V'' is the same as the element α of V iff $L=L_\alpha$ *i.e.*, iff

$$L(f)=f(\alpha) \ \forall\ f\in V'.$$

It will be in this sense that we shall regard $V''=V$.

Theorem 2. *Let V be a finite dimensional vector space over the field F. If L is a linear functional on the dual space V' of V, then there is a unique vector α in V such that $L(f)=f(\alpha)\forall f\in V'$.*

Proof. This theorem is an immediate corollary of theorem 1. We should first prove theorem 1. Then we should conclude like this :

The correspondence $\alpha\rightarrow L_\alpha$ is a one-to-one correspondence between V and V''. Therefore if $L\in V''$, then there exists a unique vector α in V such that $L=L_\alpha$ *i.e.*, such that $L\ (f)=f\ (\alpha)\forall f\in V'$.

Theorem 3. *Let V be a finite-dimensional vector space over the field F. Each basis for V' is the dual of some basis for V.*

Proof. Let $B' = \{f_1, f_2, \dots, f_n\}$ be a basis for V'. Then there exists a dual basis $(B')' = \{L_1, L_2, \dots, L_n\}$ for V'' such that

$$L_i(f_j) = \delta_{ij}. \qquad \dots(1)$$

By previous theorem, for each i there is a vector α_i in V such that

$$= L_{\alpha_i} \text{ where } L_{\alpha_i}(f) = f(\alpha_i) \ \forall\ f \in V'. \qquad \dots(2)$$

The correspondence $\alpha \leftrightarrow L_\alpha$ is an isomorphism of V onto V''. Under an isomorphism a basis is mapped onto a basis. Therefore $B = \{\alpha_1, \dots, \alpha_n\}$ is a basis for V because it is the image set of a basis for V'' under the above isomorphism.

Putting $f = f_j$ in (2), we get

$$f_j(\alpha_i) = L_{\alpha_i}(f_j) = L_i(f_j)$$
$$= \delta_{ij}. \qquad \text{[from (1)]}$$

$\therefore$ $B' = \{f_1, \dots, f_n)$ is the dual of the basis B. Hence the result.

Theorem 4. *Let V be a finite-dimensional vector space over the field F. Let B be a basis for V and B' be the dual basis of B. Then show that $B'' = (B')' = B$.*

Proof. Let $B = \{\alpha_1, \dots, \alpha_n\}$ be a basis for V,

$B' = \{f_1, \dots, f_n\}$ be the dual basis of B in V' and

$B'' = (B')' = \{L_1, \dots, L_n\}$ be the dual basis of B' in V''. Then

$$f_i(\alpha_j) = \delta_{ij},$$

and $$L_i(f_j) = \delta_{ij},\ i = 1, \dots, n\ ;\ j = 1, \dots, n.$$

If $\alpha \in V$, then there exists $L_\alpha \in V''$ such that

$$L_\alpha(f) = f(\alpha) \ \forall\ f \in V'.$$

Taking α_i in place of α, we see that for each $j = 1, \dots, n$,

$$L_{\alpha_i}(f_j) = f_j(\alpha_i) = \delta_{ij} = L_i(f_j).$$

Thus L_{α_i} and L_i agree on a basis for V'. Therefore $L_{\alpha_i} = L_i$.

If we identify V'' with V through natural isomorphism $\alpha \leftrightarrow L_\alpha$, then we consider L_α as the same element as α.

So $L_i = L_{\alpha_i} = \alpha_i$ where $i = 1, 2, \dots, n$.

Thus $B'' = B$.

Solved Examples

Ex. 1. *Let V be a vector space over the field F. Let f be a non-zero linear functional on V and let N be the null space of f. Fix a vector α_0 in V which is not in N. Prove that for each α in V there* *is a scalar* c *and a vector* β *in* N *such that* $\alpha=c\alpha_0+\beta$. *Prove that* c *and* β *are unique.*

Solution. Since f is a non-zero linear functional on V, therefore there exists a non-zero vector α_0 in V such that $f(\alpha_0)\neq 0$. Consequently $\alpha_0 \notin N$. Let $f(\alpha_0)=y\neq 0$.

Let α be any element of V and let $f(\alpha)=x$.

We have $f(\alpha)=x$

$\Rightarrow f(\alpha)=(xy^{-1})\,y$ $\quad[\because\ 0\neq y\in F\Rightarrow y^{-1}$ exists]

$\Rightarrow f(\alpha)=cy$ where $c=xy^{-1}\in F$

$\Rightarrow f(\alpha)=cf(\alpha_0)$

$\Rightarrow f(\alpha)=f(c\alpha_0)$ $\quad[\because\ f$ is a linear functional]

$\Rightarrow f(\alpha)-f(c\alpha_0)=0\Rightarrow f(\alpha-c\alpha_0)=0\Rightarrow \alpha-c\alpha_0\in N$

$\Rightarrow \alpha-c\alpha_0=\beta$ for some $\beta\in N\Rightarrow \alpha=c\alpha_0+\beta$.

If possible, let $\alpha=c'\alpha_0+\beta'$ where $c'\in F$ and $\beta'\in N$.

Then $c\alpha_0+\beta=c'\alpha_0+\beta'$ $\quad\ldots(1)$

$\Rightarrow (c-c')\,\alpha_0+(\beta-\beta')=0\Rightarrow f[(c-c')\,\alpha_0+(\beta-\beta')]=f(0)$

$\Rightarrow (c-c')\,f(\alpha_0)+f(\beta-\beta')=0$

$\Rightarrow (c-c')\,f(\alpha_0)=0$ $\quad[\because\ \beta,\beta'\in N\Rightarrow\beta-\beta'\in N\Rightarrow f(\beta-\beta')=0]$

$\Rightarrow (c-c')=0$ $\quad[\because\ f(\alpha_0)$ is a non-zero element of $F]$

$\Rightarrow c=c'$.

Putting $c=c'$ in (1), we get $c\alpha_0+\beta=c\alpha_0+\beta'\Rightarrow\beta=\beta'$.

Hence c and β are unique.

Ex. 2. *If* f *and* g *are in* V' *such that* $f(\alpha)=0\Rightarrow g(\alpha)=0$, *prove that* $g=kf$ *for some* $k\in F$.

Solution. It is given that $f(\alpha)=0\Rightarrow g(\alpha)=0$. Therefore if α belongs to the null space of f, then α also belongs to the null space of g. Thus the null space of f is a subset of the null space of g.

(i) If f is zero linear functional, then the null space of f is equal to V. Therefore in this case V is a subset of the null space of g. Hence the null space of g is equal to V. So g is also zero linear functional. Hence we have $g=kf\ \forall\ k\in F$.

(ii) Let f be a non-zero linear functional on V. Then there exists a non-zero vector $\alpha_0\in V$ such that $f(\alpha_0)=y$ where y is a non-zero element of F. Let $k=\dfrac{g(\alpha_0)}{f(\alpha_0)}$. If $\alpha\in V$, then we can write $\alpha=c\alpha_0+\beta$ where $c\in F$ and $\beta\in$ the null space of f.

We have $g(\alpha)=g(c\alpha_0+\beta)=cg(\alpha_0)+g(\beta)$

$=cg(\alpha_0)$ $\quad[\because\ \beta\in$ the null space of $f\Rightarrow f(\beta)=0$ and so $g(\beta)=0]$

Also $(kf)(\alpha)=kf(\alpha)=kf(c\alpha_0+\beta)=k\,[cf(\alpha_0)+f(\beta)]$
$=kc\,f(\alpha_0)$ $\qquad [\because\ f(\beta)=0]$
$=\dfrac{g(\alpha_0)}{f(\alpha_0)}\,cf(\alpha_0)=cg(\alpha_0)$.

Thus $g(\alpha)=(kf)(\alpha)\ \ \forall\ \alpha\in V$. Therefore $g=kf$.

§ 5. Annihilators.

Definition. *If V is a vector space over the field F and S is a subset of V, the* **annihilator** *of S is the set* S^0 *of all linear functionals f on V such that* $f(\alpha)=0\ \forall\ \alpha\in S$.

Sometimes $A(S)$ is also used to denote the annihilator of S. Thus $\qquad S^0=\{f\in V' : f(\alpha)=0\ \forall\ \alpha\in S\}$.

If $S=$ the zero subspace of V, then $S^0=V'$.

If $S=V$, then $V^0=$ the zero subspace of V'.

If V is finite dimensional and S contains a non-zero vector, then $S^0\neq V'$. If $0\neq\alpha\in S$, then there is a linear functional f on V such that $f(\alpha)\neq 0$. Thus there is $f\in V'$ such that $f\notin S^0$. Therefore $S^0\neq V'$.

Theorem 1. *If S is any subset of a vector space V(F), then* S^0 *is a subspace of V'.*

Proof. First we see that S^0 is a non-empty subset of V' because at least $\hat{0}\in S^0$. We have $\hat{0}(\alpha)=0\ \forall\ \alpha\in S$.

Let $f, g\in S^0$. Then $f(\alpha)=0\ \forall\ \alpha\in S$,

and $\qquad g(\alpha)=0\ \forall\ \alpha\in S$.

If $a, b\in F$, then

$(af+bg)(\alpha)=(af)(\alpha)+(bg)(\alpha)=af(\alpha)+bg(\alpha)=a0+b0=0$.

$\therefore\ af+bg\in S^0$.

Thus $a, b\in F$ and $f, g\in S^0 \Rightarrow af+bg\in S^0$.

$\therefore\ S^0$ is a subspace of V'.

Dimension of annihilator.

Theorem 2. *Let V be a finite dimensional vector space over the field F, and let W be a subspace of V. Then*

$$\dim W+\dim W^0=\dim V.$$

Proof. If W is the zero subspace of V, then $W^0=V'$.

$\therefore\ \dim W^0=\dim V'=\dim V$.

Also in this case $\dim W=0$. Hence the result.

Similarly the result is obvious when $W=V$.

Let us now suppose that W is a proper subspace of V. Let $\dim V=n$, and $\dim W=m$ where $0<m<n$.

Let $B_1=\{\alpha_1,\ldots,\alpha_m\}$ be a basis for W. Since B_1 is a linearly independent subset of V also, therefore it can be extended to form a basis for V. Let $B=\{\alpha_1, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_n\}$ be a basis for V.

Let $B'=\{f_1, \ldots, f_m, f_{m+1}, \ldots, f_n\}$ be the dual basis of B. Then B' is a basis for V' such that $f_i(\alpha_j)=\delta_{ij}$.

We claim that $S=\{f_{m+1}, \ldots, f_n\}$ is a basis for W^0.

Since $S \subset B'$, therefore S is linearly independent because B' is linearly independent. So S will be a basis for W^0, if W^0 is equal to the subspace of V' spanned by S *i.e.*, if $W^0=L(S)$.

First we shall show that $W^0 \subseteq L(S)$. Let $f \in W^0$. Then $f \in V'$. So let

$$f=\sum_{i=1}^{n} x_i f_i. \qquad \ldots(1)$$

Now $f \in W^0 \Rightarrow f(\alpha)=0 \ \forall \ \alpha \in W$

$\Rightarrow f(\alpha_j)=0$ for each $j=1, \ldots, m$ [$\because$ $\alpha_1, \ldots, \alpha_m$ are in W]

$$\Rightarrow \left(\sum_{i=1}^{n} x_i f_i\right)(\alpha_j)=0 \qquad \text{[from (1)]}$$

$$\Rightarrow \sum_{i=1} x_i f_i(\alpha_j)=0 \Rightarrow \sum_{i=1}^{n} x_i \delta_{ij}=0$$

$\Rightarrow x_j=0$ for each $j=1, \ldots, m$.

Putting $x_1=0, x_2=0, \ldots, x_m=0$ in (1), we get

$f=x_{m+1} f_{m+1}+\cdots+x_n f_n$

$=$ a linear combination of elements of S.

$\therefore$ $f \in L(S)$.

Thus $f \in W^0 \Rightarrow f \in L(S)$. $\therefore$ $W^0 \subseteq L(S)$.

Now we shall show that $L(S) \subseteq W^0$.

Let $g \in L(S)$. Then g is a linear combination of $f_{m+1}, \ldots, f_n$. Let

$$g=\sum_{k=m+1}^{n} y_k f_k. \qquad \ldots(2)$$

Let $\alpha \in W$. Then α is a linear combination of $\alpha_1, \ldots, \alpha_m$. Let

$$\alpha=\sum_{j=1}^{m} c_j \alpha_j. \qquad \ldots(3)$$

We have $g(\alpha)=g\left(\sum_{j=1}^{m} c_j \alpha_j\right)$ [from (3)]

$$= \sum_{j=1}^{m} c_j \, g(\alpha_j) \qquad [\because \ g \text{ is a linear functional}]$$

$$= \sum_{j=1}^{m} c_j \left(\sum_{k=m+1}^{n} y_k f_k \right) (\alpha_j) \qquad [\text{from } (2)]$$

$$= \sum_{j=1}^{m} c_j \sum_{k=m+1}^{n} y_k f_k(\alpha_j) = \sum_{j=1}^{m} c_j \sum_{k=m+1}^{n} y_k \, \delta_{kj}$$

$$= \sum_{j=1}^{m} c_j \, 0 \qquad [\because \ \delta_{kj}=0 \text{ if } k \neq j \text{ which is so for each } k=m+1, \ldots, n \text{ and for each } j=1, \ldots, m]$$

$$= 0.$$

Thus $g(\alpha)=0 \ \forall \ \alpha \in W$. Therefore $g \in W^0$.

Thus $g \in L(S) \Rightarrow g \in W^0$. $\therefore$ $L(S) \subseteq W^0$.

Hence $W^0 = L(S)$ and S is a basis for W^0.

$\therefore$ $\dim W^0 = n - m = \dim V - \dim W$

or $\dim V = \dim W + \dim W^0$.

Corollary. *If V is finite dimensional and W is a subspace of V, then W' is isomorphic to V'/W^0.*

Proof. Let $\dim V = n$ and $\dim W = m$. W' is dual space of W. So $\dim W' = \dim W = m$.

Now $\dim V'/W^0 = \dim V' - \dim W^0$

$= \dim V - (\dim V - \dim W) = \dim W = m.$

Since $\dim W' = \dim V'/W^0$, therefore $W' \cong V'/W^0$.

Annihilator of an annihilator. Let V be a vector space over the field F. If S is any subset of V, then S^0 is a subspace of V'. By the definition of an annihilator, we have

$$(S^0)^0 = S^{00} = \{L \in V'' : L(f) = 0 \ \forall \ f \in S^0\}.$$

Obviously S^{00} is a subspace of V''. But if V is finite dimensional, then we have identified V'' with V through the natural isomorphism $\alpha \to L_\alpha$. Therefore we may regard S^{00} as a subspace of V. Thus $S^{00} = \{\alpha \in V : f(\alpha) = 0 \ \forall \ f \in S^0\}$.

Theorem 3. *Let V be a finite dimensional vector space over the field F and W be a subspace of V. Then $W^{00} = W$.*

Proof. We have

$$W^0 = \{f \in V' : f(\alpha) = 0 \ \forall \ \alpha \in W\} \qquad \ldots(1)$$

and

$$W^{00} = \{\alpha \in V : f(\alpha) = 0 \ \forall \ f \in W^0\}. \qquad \ldots(2)$$

Let $\alpha \in W$. Then from (1), $f(\alpha) = 0 \ \forall \ f \in W^0$ and so from (2), $\alpha \in W^{00}$. Therefore $\alpha \in W \Rightarrow \alpha \in W^{00}$.

Thus $W \subseteq W^{00}$. Now W is a subspace of V and W^{00} is also a subspace of V. Since $W \subseteq W^{00}$, therefore W is a subspace of W^{00}.

Now $\dim W + \dim W^0 = \dim V$. [by theorem (2)]

Applying the same theorem for the vector space V' and its subspace W^0 we get

$$\dim W^0 + \dim W^{00} = \dim V' = \dim V.$$

$$\therefore \quad \dim W = \dim V - \dim W^0 = \dim V - [\dim V - \dim W^{00}]$$
$$= \dim W^{00}.$$

Since W is a subspace of W^{00} and $\dim W = \dim W^{00}$, therefore $W = W^{00}$.

Solved Examples

Ex. 1. *If S_1 and S_2 are two subsets of a vector space V such that $S_1 \subseteq S_2$, then show that $S_2^0 \subseteq S_1^0$.*

Solution. Let $f \in S_2^0$. Then

$$f(\alpha) = 0 \ \forall \ \alpha \in S_2$$
$$\Rightarrow f(\alpha) = 0 \ \forall \ \alpha \in S_1 \qquad [\because \ S_1 \subseteq S_2]$$
$$\Rightarrow f \in S_1^0.$$

$\therefore \quad S_2^0 \subseteq S_1^0$.

Ex. 2. *Let V be a vector space over the field F. If S is any subset of V, then show that $S^0 = [L(S)]^0$.*

Solution. We know that $S \subseteq L(S)$.

$\therefore \quad [L(S)]^0 \subseteq S^0$.

Now let $f \in S^0$. Then $f(\alpha) = 0 \ \forall \ \alpha \in S$.

If β is any element of $L(S)$, then

$$\beta = \sum_{i=1}^{n} x_i \alpha_i \text{ where each } \alpha_i \in S.$$

We have $f(\beta) = \sum_{i=1}^{n} x_i f(\alpha_i) = 0$, since each $f(\alpha_i) = 0$.

Thus $f(\beta) = 0 \ \forall \ \beta \in L(S)$.

$\therefore \quad f \in (L(S))^0$.

Therefore $\quad S^0 \subseteq (L(S))^0$. ...(2)

From (1) and (2), we conclude that $S^0 = (L(S))^0$.

Ex. 3. *Let V be a finite-dimensional vector space over the field F. If S is any subset of V, then $S^{00} = L(S)$.*

Solution. We have $S^0 = (L(S))^0$. [See Ex. 2]

$\therefore \quad S^{00} = (L(S))^{00}$. ...(1)

But V is finite dimensional and $L(S)$ is a subspace of V. Therefore by theorem 3, $(L(S))^{00}=L(S)$.

. from (1), we have $S^{00}=L(S)$.

Ex. 4. *Let V be a finite dimensional vector space over the field F. If W_1 and W_2 are subspaces of V, then $W_1^0=W_2^0$ iff $W_1=W_2$.*

Solution. We have $W_1=W_2 \Rightarrow W_1^0=W_2^0$.

Conversely, let $W_1^0=W_2^0$. Then $W_1^{00}=W_2^{00} \Rightarrow W_1=W_2$.

Ex. 5. *Let W_1 and W_2 be subspaces of a finite dimensional vector space V.*

(*a*) *Prove that* $(W_1+W_2)^0=W_1^0 \cap W_2^0$.

(*b*) *Prove that* $(W_1 \cap W_2)^0=W_1^0+W_2^0$.

Solution. (a) First we shall prove that

$$W_1^0 \cap W_2^0 \subseteq (W_1+W_2)^0.$$

Let $f \in W_1^0 \cap W_2^0$. Then $f \in W_1^0, f \in W_2^0$.

Suppose α is any vector in W_1+W_2. Then

$$\alpha=\alpha_1+\alpha_2 \text{ where } \alpha_1 \in W_1, \alpha_2 \in W_2.$$

We have $f(\alpha)=f(\alpha_1+\alpha_2)=f(\alpha_1)+f(\alpha_2)$

$=0+0$ [$\because$ $\alpha_1 \in W_1$ and $f \in W_1^0 \Rightarrow f(\alpha_1)=0$ and similarly $f(\alpha_2)=0$]

$=0.$

Thus $f(\alpha)=0 \;\forall\; \alpha \in W_1+W_2$.

$\therefore$ $f \in (W_1+W_2)^0$.

$\therefore$ $W_1^0 \cap W_2^0 \subseteq (W_1+W_2)^0$. ...(1)

Now we shall prove that $(W_1+W_2)^0 \subseteq W_1^0 \cap W_2^0$.

We have $W_1 \subseteq W_1+W_2$.

$\therefore$ $(W_1+W_2)^0 \subseteq W_1^0$ ...(2)

Similarly $W_2 \subseteq W_1+W_2$.

$\therefore$ $(W_1+W_2)^0 \subseteq W_2^0$. ...(3)

From (2) and (3), we have

$(W_1+W_2)^0 \subseteq W_1^0 \cap W_2^0$. ...(4)

From (1) and (4), we have $(W_1+W_2)^0=W_1^0 \cap W_2^0$.

(b) Let us use the result (a) for the vector space V' in place of the vector space V. Thus replacing W_1 be W_1^0 and W_2 by W_2^0 in (a) we get

$$(W_1^0+W_2^0)^0=W_1^{00} \cap W_2^{00}$$

$\Rightarrow (W_1^0+W_2^0)^0=W_1 \cap W_2$ [$\because$ $W_1^{00}=W_1$ etc.]

$\Rightarrow (W_1^0+W_2^0)^{00}=(W_1 \cap W_2)^0 \Rightarrow W_1^0+W_2^0=(W_1 \cap W_2)^0$.

8

Modules

The concept of module is a generalisation of that of a vector space. In a vector space the scalars are elements of a field while in a module we shall allow the scalars to be elements of an arbitrary ring.

§ 1. Module. Definition.

A non-empty set M is said to be a left module over a ring R (or, a left R-module) if M is an abelian group under an operation $+$ such that for every $r \in R$, $m \in M$ there exists a unique element $rm \in M$ subject to the conditions :

(1) $r(a+b)=ra+rb$

(2) $(r+s)a=ra+sa$

(3) $r(sa)=(rs)a$

for all $a, b \in M$ and $r, s \in R$.

If R is a ring with unity 1, then a left R-module is said to be *unital* if $1m=m$ for all $m \in M$. If the ring R is a field, then a unital left R-module is nothing but a vector space over the field R.

We have called the algebraic structure defined above a left R-module because we have allowed multiplication by the elements of R from the left. In a similar fashion we can define a *right R-module* by modifying the conditions (1), (2) and (3) in the above definition in the following manner :

(1) $(a+b)r=ar+br$

(2) $a(r+s)=ar+as$

(3) $(ar)s=a(rs)$

for all $a, b \in M$ and $r, s \in R$.

It should be noted that the distinction between a left R-module and a right R-module is merely that of notation. The theory of right R-modules can be developed in the same manner as the theory of left R-modules. We shall develop here the theory of left R-modules. If M is a left R-module, then we shall omit the repeated use of the adjective *left* and we shall simply call it a *R-module.*

Some Examples of Modules

Example 1. *Every abelian group G is a module over the ring of integers* **I.**

Let G be an abelian group, the operation in G being denoted by $+$ and the identity element of G by 0. For any integer n and for any element a of G we define na in the following manner :

If n is a positive integer, we define $na=a+a+a+\ldots$ upto n terms. If $n=0$, we define $0a=0$ where 0 on the right hand side is the identity of G. If n is a negative integer, say $n=-m$ where m is a positive integer, we define $(-m)\,a=-(ma)$, where $-(ma)$ denotes the inverse of ma in G. It can by easily seen that

$$-(ma)=m\,(-a).$$

Now G will be a module over the ring of integers **I**, if we prove that

(1) $m\,(a+b)=ma+mb$

(2) $(m+n)\,a=ma+na$

(3) $m\,(na)=(mn)\,a$

for all $m, n \in \mathbf{I}$ and $a, b \in G$. We shall prove these one by one.

(1) Let us prove that $m(a+b)=ma+mb \ \forall \ m \in \mathbf{I}$, and $a, b \in G$.

If m is a positive integer then

$$\begin{aligned} m\,(a+b)&=(a+b)+(a+b)+(a+b)+\ldots \text{upto } m \text{ terms} \\ &=(a+a+a+\ldots \text{upto } m \text{ terms})+(b+b+\ldots \text{upto } m \text{ terms}) \\ &=ma+mb. \end{aligned}$$

If $m=0$, then $0\,(a+b)=0=0+0=0a+0b$.

If m is a negative integer, say $m=-p$ where p is a positive integer, then

$$\begin{aligned} m\,(a+b)&=(-p)\,(a+b)=-[p\,(a+b)]=-(pa+pb) \\ &=-(pa)+[-(pb)]=(-p)\,a+(-p)\,b=ma+mb. \end{aligned}$$

Thus (1) is true. The proofs of (2) and (3) have been left for the reader. Hence G is a module over the ring of integers.

Example 2. *Let R be a ring and let M be a left ideal of R. For $r \in R$, $m \in M$ let rm be the product of these elements as elements in R. Then M is an R-module.*

Since M is a left ideal of R, therefore M is an additive abelian group. Also if $r \in R$ and $m \in M$, then $rm \in M$ because M is a left ideal of R. Further

(1) $r(m_1+m_2)=rm_1+rm_2$ for all $r\in R$ and $m_1, m_2\in M$. This result is a consequence of left distributive law in R.

(2) $(r_1+r_2)m=r_1m+r_2m$ for all $r_1, r_2\in R$ and $m\in M$. This result is a consequence of right distributive law in R.

(3) $r(sm)=(rs)\,m$ for all $r, s\in R$ and $m\in M$. This follows from associativity.

Hence M is an R-module.

Example 3. *Every ring R is an R-module over itself.*

This result follows as a special case from example 2 if we take $M=R$.

Example 4. *Let R be any ring and λ be a left ideal of R. Let M consist of all the cosets, $a+\lambda$, where $a\in R$, of λ in R. Thus $M=\{a+\lambda : a\in R\}$. It can be shown that M is an R-module if the two requisite compositions are defined as follows:*

$(a+\lambda)+(b+\lambda)=(a+b)+\lambda$ [*Addition of elements of M*]

$r\,(a+\lambda)=ra+\lambda$ [*Multiplication of an element of M by an element of R*]

Since $(a+b)+\lambda$ and $ra+\lambda$ are also cosets of λ in R, therefore M is closed with respect to the two compositions. First of all we shall show that both these compositions are well defined.

Let $a+\lambda=a'+\lambda,\ a, a'\in R$

and $b+\lambda=b'+\lambda,\ b, b'\in R$.

We have $a+\lambda=a'+\lambda \Rightarrow a-a'\in\lambda$

and $b+\lambda=b'+\lambda \Rightarrow b-b'\in\lambda$.

Now λ is a left ideal, therefore $a-a'\in\lambda,\ b-b'\in\lambda$

$\Rightarrow (a-a')+(b-b')\in\lambda \Rightarrow (a+b)-(a'+b')\in\lambda$

$\Rightarrow (a+b)+\lambda=(a'+b')+\lambda \Rightarrow (a+\lambda)+(b+\lambda)=(a'+\lambda)+(b'+\lambda)$.

Therefore addition in M is well defined.

Again λ is a left ideal. Therefore $r\in R,\ a-a'\in\lambda$

$\Rightarrow r\,(a-a')\in\lambda \Rightarrow ra-ra'\in\lambda \Rightarrow ra+\lambda=ra'+\lambda$

$\Rightarrow r\,(a+\lambda)=r\,(a'+\lambda)$.

Therefore the multiplication of the elements of M by the elements of R is also well defined.

Commutativity of addition in M. We have

$(a+\lambda)+(b+\lambda)=(a+b)+\lambda=(b+a)+\lambda=(b+\lambda)+(a+\lambda)$.

Associativity of addition in M. We have

$(a+\lambda)+[(b+\lambda)+(c+\lambda)]=(a+\lambda)+[(b+c)+\lambda]=[a+(b+c)]+\lambda$

$=[(a+b)+c]+\lambda=[(a+b)+\lambda]+(c+\lambda)=[(a+\lambda)+(b+\lambda)]+(c+\lambda)$.

The coset $0+\lambda$ is the additive identity in M for we have

$$(0+\lambda)+(a+\lambda)=(0+a)+\lambda=a+\lambda.$$

The coset $(-a)+\lambda$ is the additive inverse of the element $a+\lambda$ of M. We have $(-a+\lambda)+(a+\lambda)=(-a+a)+\lambda=0+\lambda=$ the identity element of M.

Thus M is an additive abelian group. We further observe that

(1) $r[(a+\lambda)+(b+\lambda)]=r[(a+b)+\lambda]=r(a+b)+\lambda$
$=(ra+rb)+\lambda=(ra+\lambda)+(rb+\lambda)=r(a+\lambda)+r(b+\lambda)$.

(2) $(r+s)(a+\lambda)=[(r+s)a]+\lambda=(ra+sa)+\lambda$
$=(ra+\lambda)+(sa+\lambda)=r(a+\lambda)+s(a+\lambda)$.

(3) $r[s(a+\lambda)]=r(sa+\lambda)=[r(sa)]+\lambda=[(rs)a]+\lambda$
$=(rs)(a+\lambda)$.

Hence M is an R-module. M is usually written as $R-\lambda$ (or, sometimes R/λ) and is called the *difference* (or *quotient*) *module of R by* λ.

Example 5. *The module of all ordered n-tuples of elements of a ring R.*

Let $M=\{(a_1, a_2, \ldots, a_n) : a_1, a_2, \ldots, a_n \in R\}$. *If we define addition in M by* $(a_1, a_2, \ldots, a_n)+(b_1, b_2, \ldots, b_n)=(a_1+b_1, a_2+b_2, \ldots, a_n+b_n)$ *and multiplication of an element* $(a_1, a_2, \ldots, a_n)$ *of M by an element r of R by* $r(a_1, a_2, \ldots, a_n)=(ra_1, ra_2, \ldots, ra_n)$, *then it can be easily shown that M is an R-module.*

§ 2. General properties of modules.

Theorem. *Let M be an R-module. Then*

(i) $r0=0 \ \forall r\in R$. (ii) $0a=0 \ \forall a\in M$.

(iii) $(-r)a=-(ra)=r(-a)$ *for all* $r\in R, a\in M$.

(iv) $(-r)(-a)=ra$ *for all* $r\in R, a\in M$.

(v) $r(a-b)=ra-rb$ *for all* $r\in R$, $a, b\in M$.

(vi) $(r-s)a=ra-sa$ *for all* $r, s\in R$, $a\in M$.

Proof. (i) We have

$$r0=r(0+0) \quad [\because\ 0=0+0 \text{ in the group } M]$$
$$=r0+r0.$$

$\therefore\ 0+r0=r0+r0.$ $[\because\ r0\in M \text{ and } 0+r0=r0]$

Now M is an abelian group with respect to addition. Therefore by right cancellation law in M, we get $0=r0$.

The proofs of the remaining parts have been left for the reader. Proceed as in theorem 1, page 402, chapter 6 (on vector spaces).

§ 3. Submodules. Definition.

A non-empty subset S of an R-module M is said to be an R-submodule of M (or, simply a submodule of M) if (i) S is an addi-

tive subgroup of M and (ii) $r \in R$, $a \in S \Rightarrow ra \in S$. If M is any R-module, then M itself and (0) *i.e.*, the subset of M consisting of the identity element of M alone are always submodules of M. These two are called *improper submodules.* If M has any other submodule, then it is called a *proper submodule.*

Irreducible R-module. Definition. *An R-module M is said to be irreducible if its only submodules are* (0) *and M.*

Intersection of submodules

Theorem 1. *If A and B are two submodules of an R-module M, then* $A \cap B$ *is also a submodule of M.*

Proof. Since A and B are additive subgroups of M, therefore $A \cap B$ is also an additive subgroup of M.

Now let $r \in R$ and $a \in A \cap B$. Then $a \in A$ and $a \in B$. Since A is a submodule of M, therefore $r \in R$, $a \in A \Rightarrow ra \in A$. Also B is a submodule of M. Therefore $r \in R$, $a \in B \Rightarrow ra \in B$. Thus $ra \in A \cap B$. Hence $A \cap B$ is a submodule of M.

Theorem 2 *Arbitrary intersection of submodules i.e., the intersection of any family of submodules of a module is a submodule.*

The proof of this theorem has been left for the reader. Proceed as in theorem 3 page 409, chapter 6 (on vector spaces).

Submodule generated by a subset of a module. Let M be an R-module and S be a non-empty subset of M. If A is a submodule of M containing S and is itself contained in every submodule of M containing S, then A is called the *submodule of M generated by S.* The submodule of M generated by S will be denoted by the symbol (S). It should be noted that (S) is the smallest submodule of M containing S. It can be easily seen that the intersection of all the submodules of M containing S is the submodule of M generated by S.

Theorem 3. *Show that the submodule of a unital R-module M generated by a subset S of M consists of all linear combinations of elements in S.*

Proof. Let $L(S)$ denote the set of all linear combinations of the elements of S *i.e.*, let $L(S) = \{r_1a_1 + r_2a_2 + \dots + r_na_n : a_1, a_2, \dots, a_n$ is any arbitrary finite subset of S and $r_1, r_2, \dots, r_n$ is any arbitrary finite subset of the ring$\}$.

First we shall show that $L(S)$ is a submodule of M. Let $a = r_1a_1 + r_2a_2 + \dots + r_na_n$, $b = s_1b_1 + s_2b_2 + \dots + s_mb_m$ be any two elements of $L(S)$. Here the r_i's and the s_i's are elements of R and the a_i's and b_i's are elements of S. We have $a \quad b = r_1a_1 + r_2a_2 + \dots +$

$r_n a_n + (-s_1) b_1 + (-s_2) b_2 + \ldots + (-s_m) b_m$. Obviously $a-b$ is an element of $L(S)$ because it is also a linear combination of some elements of S. Thus $a, b \in L(S) \Rightarrow a-b \in L(S)$. Hence $L(S)$ is an additive subgroup of M.

Now if r is any element of R and $a = r_1 a_1 + r_2 a_2 + \ldots + r_n a_n$ is any element of $L(S)$, then $ra = r(r_1 a_1 + r_2 a_2 + \ldots + r_n a_n) = r(r_1 a_1) + \ldots + r(r_n a_n) = (rr_1) a_1 + \ldots + (rr_n) a_n$. Since $rr_1, \ldots, rr_n \in R$, therefore ra is also a linear combination of some elements of S. Thus $r \in R$, $a \in L(S) \Rightarrow ra \in L(S)$. Therefore $L(S)$ is a submodule of M.

Also each element of S belongs to $L(S)$ because if $a_i \in S$, then $a_i = 1a_i$ where 1 is the unity element of the ring R. Note that M is a unital R-module. Now $a_i = 1a_i \Rightarrow a_i \in L(S)$. Thus $L(S)$ is a submodule of M and S is contained in $L(S)$.

Now if W is any submodule of M containing S, then each element of $L(S)$ must be in W because W is to be closed under scalar multiplication and addition Therefore $L(S)$ will be contained in W.

Hence $L(S) = (S)$ *i e.*, $L(S)$ is the submodule of M generated by S.

§ 4. **Linear sum of two submodules. Definition.** Let A and B be two submodules of an R-module M. Then the linear sum of the submodules A and B denoted by $A+B$ is the set of sums $a+b$ such that $a \in A$, $b \in B$.

Thus $A+B = \{a+b : a \in A, b \in B\}$.

Theorem. *If A and B are submodules of an R-module M, then $A+B$ is also a submodule of M.*

Proof. Let $c = a_1 + b_1$, $d = a_2 + b_2$ be any two elements of $A+B$. Then $a_1, a_2 \in A$ and $b_1, b_2 \in B$ We have $c-d = (a_1+b_1) - (a_2+b_2) = (a_1 - a_2) + (b_1 - b_2)$. Since A is an additive subgroup of M, therefore $a_1, a_2 \in A \Rightarrow a_1 - a_2 \in A$. Similarly $b_1 - b_2 \in B$. Therefore $(a_1 - a_2) + (b_1 - b_2) \in A+B$. Thus $c, d \in A+B \Rightarrow c-d \in A+B$. Therefore $A+B$ is an additive subgroup of M.

Now let $r \in R$ and $c = a_1 + b_1 \in A+B$. We have $rc = r(a_1 + b_1) = ra_1 + rb_1$. Since A is a submodule of M, therefore $r \in R$, $a_1 \in A \Rightarrow ra_1 \in A$. Similarly $rb_1 \in B$. Thus $ra_1 + rb_1 \in A+B$. In this way $r \in R$, $c \in A+B \Rightarrow rc \in A+B$. Hence $A+B$ is a submodule of M.

§ 5. **Direct sum of submodules. Module as a direct sum of submodules. Definition.** *Let M be an R-module and $M_1, M_2, \ldots, M_n$*

be submodules of M. *Then* M *is said to be the direct sum of* $M_1, M_2, \ldots, M_n$ *if every element* $a \in M$ *can be written in one and only one way as* $a = a_1 + a_2 + \ldots + a_n$ *where* $a_1 \in M_1, a_2 \in M_2, \ldots, a_n \in M_n$.

If a module M is a direct sum of its submodules M_1 and M_2 then we should have not only $M = M_1 + M_2$ but also that each element of M can be uniquely expressed as sum of an element of M_1 and an element of M_2. Symbolically the direct sum is represented by the notation $M = M_1 \oplus M_2$.

Theorem. *The necessary and sufficient conditions for a module* M *to be a direct sum of its two submodules* M_1 *and* M_2 *are that*

(*i*) $M = M_1 + M_2$

and (*ii*) $M_1 \cap M_2 = \{0\}$.

The proof of this theorem has been left for the reader.

§ 6. Homomorphism of modules or linear transformations.

Definition. *Let* M *and* N *be two* R*-modules. A mapping* T *from* M *into* N *is called a homomorphism (or* R*-homomorphism or module homomorphism) if*

(*i*) $T(m_1 + m_2) = T(m_1) + T(m_2) \ \forall \ m_1, m_2 \in M$

(*ii*) $T(rm) = rT(m) \ \forall \ r \in R, m \in M$.

If T is a homomorphism of M onto N, then N is called a homomorphic image of M.

If T is a homomorphism of M into N and if the mapping T is one-to-one, then T is called an *isomorphism of* M into N.

Note. Sometimes the image of m under T *i e.*, Tm or $T(m)$ is also written as mT.

Theorem 1. *If* T *is a homomorphism of an* R*-module* M *into an* R*-module* N, *then*

(*i*) $T(0) = 0$,

(*ii*) $T(-m) = -T(m)$,

(*iii*) $T(m_1 - m_2) = T(m_1) - T(m_2)$ *for all* $m, m_1, m_2 \in M$.

The proof of this theorem has been left for the reader.

Kernel of a homomorphism. Definition. *Let* T *be a homomorphism of an* R*-module* M *into an* R*-module* N. *The kernel* $K(T)$ *of* T *is defined as*

$K(T) = \{m \in M : T(m) = 0$ *where* 0 *is the identity element of the additive group* $N\}$.

Theorem 2. *The kernel of a homomorphism is a submodule.*

Proof. Let $K(T)$ be the kernel of the homomorphism T of an R-module M into an R-module N. Then $K(T) = \{m \in M :$

$T(m)=0\}$. To prove that $K(T)$ is a submodule of M. Since $T(0)=0$, therefore at least $0 \in K(T)$. Thus $K(T)$ is a non-empty subset of M.

Now let $m_1, m_2 \in K(T)$. Then $T(m_1)=0$ and $T(m_2)=0$. We have $T(m_1-m_2)=T(m_1)-T(m_2)=0-0=0$. Therefore $m_1-m_2 \in K(T)$. Thus $K(T)$ is an additive subgroup of M.

Again let $r \in R$ and $m \in K(T)$. Then $T(m)=0$. We have $T(rm)=r\,T(m)=r0=0$. Therefore $rm \in K(T)$. Hence $K(T)$ is a submodule of M.

Theorem 3. *The range of a homomorphism is a submodule.*

Proof. Let T be a homomorphism of an R-module M into an R-module N. Let $I(T)$ denote the range of T. Then $I(T)=\{T(m) : m \in M\}$. To prove that $I(T)$ is a submodule of N.

Let $T(m_1)$, $T(m_2)$ be any two elements of $I(T)$ where $m_1, m_2 \in M$. We have $T(m_1)-T(m_2)=T(m_1-m_2) \in I(T)$ since $m_1-m_2 \in M$. Therefore $I(T)$ is an additive subgroup of N.

Again let r be any element of R and $T(m)$ be any element of $I(T)$ where $m \in M$. We have $r\,T(m)=T(rm) \in I(T)$ since $rm \in M$. Hence $I(T)$ is a submodule of N.

Theorem 4. *Let T be a module homomorphism. Show that T is an isomorphism if and only if $K(T)=(0)$.*

Proof. Let T be a homomorphism of an R-module M into an R-module N. First we shall prove that T is an isomorphism if kernel of T *i.e.*, $K(T)=(0)$. If $m_1, m_2 \in M$, then

$$T(m_1)=T(m_2) \Rightarrow T(m_1)-T(m_2)=0 \Rightarrow T(m_1-m_2)=0$$
$$\Rightarrow m_1-m_2 \in K(T) \Rightarrow m_1-m_2=0, \text{ since } K(T)=(0).$$
$$\Rightarrow m_1=m_2 \Rightarrow T \text{ is one-to-one} \Rightarrow T \text{ is an isomorphism.}$$

Conversely suppose that T is an isomorphism *i.e.*, T is one-to-one. Then to show that $K(T)=(0)$. Let $m \in K(T)$. Then $T(m)=0=T(0)$. But T is one-to-one. Therefore $T(m)=T(0)$ gives $m=0$. Thus $m \in K(T) \Rightarrow m=0$. Hence $K(T)=(0)$.

§ 7. **Quotient modules.** Let A be any submodule of an R-module M. Then A is a subgroup of the additive group M. If $m \in M$, then $A+m$ is a coset of A in M. Let M/A denote the set of all cosets of A in M *i.e.*, let $M/A=\{A+m : m \in M\}$. We know that if $A+m_1$ and $A+m_2$ are two cosets of A in M, then $A+m_1=A+m_2 \Leftrightarrow m_1-m_2 \in A$. Now we shall give a module structure to the set M/A over the same ring R. For this we shall have to define addition in M/A and multiplication of an element of M/A by an element of R *i.e.*, scalar multiplication.

Theorem 1. *If A is any submodule of an R-module M, then the set M/A of all cosets A+m where m is any arbitrary element of M, is a module over R for the addition and scalar multiplication compositions defined as follows :*

$(A+m_1)+(A+m_2)=A+(m_1+m_2) \ \forall \ m_1, m_2 \in M$

$r(A+m)=A+rm \ \forall \ r\in R, m\in M.$

The R-module M/A is called the **Quotient module** of M relative to the submodule A.

For proof of this theorem proceed as in theorem on page 441 in the chapter 6 on vector spaces.

Theorem 2. *Suppose M is an R-module and A is a submodule of M. Let T be mapping from M into M/A defined by* $T(m)=A+m$ $\forall \ m \in M$. *Then T is an R-homomorphism of M onto M/A and kernel* $T=A$.

Theorem 3. Fundamental theorem on homomorphism of modules. *If T is a homomorphism of an R-module M onto an R-module N with* $K(T)=A$, *then N is isomorphic (as a module) to M/A.*

The proofs of both these theorems have been left for the reader.

§ 8. **Cyclic modules. Definition.** *An R-module M is said to by cyclic if there is an element* $m_0 \in M$ *such that every* $m\in M$ *is of the form,* $m=rm_0$ *where* $r\in R$. *Also* m_0 *is called a generator of M and we write* $M=(m_0)$.

For R, the ring of integers, a cyclic R-module is nothing more than a cyclic group.

Theorem. *Let M be a unital R-module and for a fixed element* $m\in M$ *let* $A=\{rm : r\in R\}$. *Then A is a cyclic submodule of M generated by m.*

The proof of this theorem has been left for the reader.

§ 9. **Finitely generated Modules. Definition.** *An R-module M is said to be finitely generated if there exist elements* $a_1, a_2, \ldots, a_n \in M$ *such that every m in M is of the form* $m=r_1a_1+r_2a_2+\ldots+r_na_n$ *where* $r_1, r_2, \ldots, r_n \in R$.

Now we come to the main theorem of this chapter. This theorem is known as the *fundamental theorem on finitely generated untial modules over Euclidean rings.*

Theorem. *Let R be a Euclidean ring; then any finitely generated R module, M, is the direct sum of a finite number of cyclic modules.*

Proof. It is given that M is a finitely generated R-module where R is a Euclidean ring. We shall call those generating sets which have as few elements as possible ***minimal generating sets*** and the number of elements in such a minimal generating set the ***rank*** of M. We shall prove the theorem by induction on the rank of M.

To start the induction we see that if M is of rank 1, then M is generated by a single element and so M is cyclic and the theorem is true. Now assume as our induction hypothesis that the theorem is true for all R-modules of rank $k-1$. If M is a R-module of rank k, we are to show that the theorem is true for M.

If for any given minimal generating set $c_1, c_2, \dots, c_k$ of M, any relation of the form $r_1c_1+r_2c_2+\dots+r_kc_k=0$ $(r_i \in R)$ implies that $r_1c_1=r_2c_2=\dots=r_kc_k=0$, then obviously M is the direct sum of $M_1, M_2, \dots, M_k$ where each M_i is a cyclic submodule of M generated by c_i. So in this case we are left nothing to prove. Consequently, given any minimal generating set $b_1, b_2, \dots, b_k$ of M, there must be elements $r_1, r_2, \dots, r_k$ of R such that $r_1b_1+r_2b_2+\dots+r_kb_k=0$ and in which not all of $r_1b_1, r_2b_2, \dots, r_kb_k$ are 0. Among all possible relations of this type for all minimal generating sets let s_1 be the element of R whose d-value $d(s_1)$ is minimal. Let the generating set for which it occurs be $a_1, a_2, \dots, a_k$. Thus

$$s_1a_1+s_2a_2+\dots+s_ka_k=0. \qquad \dots(1)$$

Now we claim that if

$$r_1a_1+r_2a_2+\dots+r_ka_k=0, \qquad \dots(2)$$

then s_1 is a divisor of r. Since s_1, r_1 are elements of a Euclidean ring R, therefore there exist $m, t \in R$ such that $r_1=ms_1+t$ where either $t=0$ or $d(t) < d(s_1)$. Multiplying (1) by m and subtracting from (2), we get

$$(r_1-ms_1)\,a_1+(r_2-ms_2)\,a_2+\dots+(r_k-ms_k)\,a_k=0$$

or
$$ta_1+(r_2-ms_2)\,a_2+\dots+(r_k-ms_k)\,a_k=0. \qquad \dots(3)$$

If $t \neq 0$, then $d(t) < d(s_1)$ and so the relation (3) contradicts our choice of s_1. Therefore we must have $t=0$. Then $r_1=ms_1+t$ gives $r_1=ms_1$ and so s_1 is a divisor of r_1.

Our next claim is that s_1 is a divisor of s_i, for $i=2, 3, \dots, k$. Let us show that $s_1 \mid s_2$. Since s_1, s_2 are elements of R, therefore there exist $m_2, t_2 \in R$ such that $s_2=m_2s_1+t_2$ where either $t_2=0$ or $d(t_2) < d(s_1)$. Now $a_1'=a_1+m_2a_2, a_2, a_3, \dots, a_k$ also generate M. We have

$$\begin{aligned} s_1a_1'+t_2a_2+s_3a_3+\dots+s_ka_k &= s_1\,(a_1+m_2a_2)+t_2a_2+s_3a_3+\dots+s_ka_k \\ &= s_1a_1+(s_1m_2+t_2)\,a_2+s_3a_3+\dots+s_ka_k \\ &= s_1a_1+s_2a_2+\dots+s_ka_k=0. \qquad \dots(4) \end{aligned}$$

If $t_2 \neq 0$, then $d(t_2) < d(s_1)$. Therefore the relation (4) *i.e.*, $s_1a_1'+t_2a_2+\dots+s_ka_k=0$ contradicts our choice of s_1. So we must have $t_2=0$ and this makes $s_2=m_2s_1$ and thus $s_1 \mid s_2$ Similarly we can show that $s_1 \mid s_i$, $i=3, 4, \dots, k$. Let us write $s_2=m_2s_1$, $s_3=m_3s_1$, $\dots$, $s_k=m_ks_1$.

Consider the set $a_1^*, a_2, a_3, \dots, a_k$ where $a_1^*=a_1+m_2a_2+\dots+m_ka_k$. Obviously $a_1^*, a_2, \dots, a_k$ generate M. Let M_1 be the cyclic submodule of M generated by a_1^* and M_2 be the submodule of M generated by $a_2, \dots, a_k$ We claim that $M=M_1 \oplus M_2$. For this we are to show that $M=M_1+M_2$ and $M_1 \cap M_2=(0)$. Since M is generated by $a_1^*, a_2, \dots, a_k$, therefore $M=M_1+M_2$. Now let $\beta \in M_1 \cap M_2$. Then $\beta \in M_1$, $\beta \in M_2$ Since $\beta \in M_1$, therefore $\beta=r_1a_1^*$ for some $r_1 \in R$. Also $\beta \in M_2 \Rightarrow \beta=r_2a_2+\dots+r_ka_k$ for some $r_2, \dots, r_k \in R$. From these we get

$$r_1a_1^*=r_2a_2+\dots+r_ka_k$$

$$\Rightarrow r_1a_1^*-r_2a_2-\dots-r_ka_k=0$$

$$\Rightarrow r_1(a_1+m_2a_2+\dots+m_ka_k)-r_2a_2-\dots-r_ka_k=0$$

$$\Rightarrow r_1a_1+(r_1m_2-r_2)a_2+\dots+(r_1m_k-r_k)a_k=0. \qquad \dots(5)$$

In the relation (5) between $a_1, \dots, a_k$ the coefficient of a_1 is r_1. Therefore by what we have proved above s_1 is a divisor of r_1. Let $r_1=ps_1$ where $p \in R$. We have

$$\begin{aligned}\beta&=r_1a_1^*=(ps_1)a_1^*=p(s_1a_1^*)\\&=p[s_1(a_1+m_2a_2+\dots+m_ka_k)]\\&=p[s_1a_1+s_1m_2a_2+\dots+s_1m_ka_k]\\&=p(s_1a_1+s_2a_2+\dots+s_ka_k)=p\,0=0.\end{aligned}$$

Thus $\beta \in M_1 \cap M_2 \Rightarrow \beta=0$ Therefore $M_1 \cap M_2=(0)$. Hence $M=M_1 \oplus M_2$. Now M_2 is generated by $a_2, \dots, a_k$. Therefore rank of M_2 is at most $k-1$. So by our induction hypothesis M_2 is the direct sum of cyclic modules. Hence M is the direct sum of cyclic modules. The proof of the theorem is now complete by induction.

Corollary. *Any finite abelian group is the direct product (sum) of cyclic groups.*

Proof. Let G be a finite abelian group. Then G is an R-module if R is the ring of integers [See example 1 page 469]. Obviously G is finitely generated, in fact G is generated by the finite set consisting of all its elements. Also the ring of integers is a Euclidean ring. Therefore by the above theorem G is the direct sum of cyclic submodules. But a cyclic module over the ring of integers is nothing more than a cyclic group. Hence G is the direct sum of cyclic groups.

Solved Examples

Ex. 1. *Let M be an R-module. Show that if I is a right ideal of R, then the totality of elements $y \in M$ such that $by=0$, for all $b \in I$ is a submodule of M.*

Solution. Let $A=\{y : y \in M$ and $by=0$ for all $b \in I\}$. To show that A is a submodule of M. Let $y_1, y_2 \in A$. Then $by_1=0$, $by_2=0$ for all $b \in I$. For all $b \in I$, we have $b(y_1-y_2)=by_1-by_2$ $=0-0=0$. Therefore $y_1-y_2 \in A$. Thus A is an additive subgroup of M.

Now let $r \in R$ and $y \in A$. Then $by=0$ for all $b \in I$.

If b is any element of I, then $br \in I$ because I is a right ideal of R. Now $b(ry)=(br)y=0$ because $br \in I$ and $by=0$ for all $b \in I$. Thus $b(ry)=0$ for all $b \in I$. Therefore $ry \in A$. Hence A is a submodule of M.

Ex. 2. *Prove that any unital, irreducible R-module is cyclic.*

Solution. Let M be a unital irreducible R-module. Then the only submodules of M are (0) and M itself. To prove that M is cyclic. If $M=(0)$, then obviously M is cyclic. So let us take $M \neq (0)$. Let $m_0 \in M$ and let $m_0 \neq 0$. Let $A=\{rm_0 : r \in R\}$. Then obviously A is a submodule of M. If 1 is the unity element of the ring R, then M is unital implies that $1m_0=m_0$. From our definition of A we see that $1m_0=m_0 \in A$. Since $m_0 \neq 0$, therefore A is a submodule of M such that $A \neq (0)$. But M is irreducible. Therefore we must have $A=M$. Thus $M=\{rm_0 : r \in R\}$. Therefore if m is an arbitrary element of M, then $m=rm_0$ for some $r \in R$. Thus M is cyclic.

Ex. 3. *Let M be an R-module ; if $m \in M$ let $\lambda(m)=\{x \in R : xm=0\}$. Show that $\lambda(m)$ is a left ideal of R.*

Solution. Let $x_1, x_2 \in \lambda(m)$. Then $x_1m=0$ and $x_2m=0$. We have $(x_1-x_2)m=x_1m-x_2m=0-0=0$. Therefore $x_1-x_2 \in \lambda(m)$. Thus $\lambda(m)$ is an additive subgroup of R.

Now let $r \in R$ and $x \in \lambda(m)$. Then $xm=0$. We have $(rx)m=r(xm)=r0=0$. Therefore $rx \in \lambda(m)$. Hence $\lambda(m)$ is a left ideal of R.

Ex. 4. *If λ is a left ideal of R and M is an R-module show that for $m \in M$, $\lambda m=\{xm : x \in \lambda\}$ is a submodule of M.*

Solution. Let x_1m and x_2m be any two elements of λm. Then $x_1, x_2 \in \lambda$. We have $x_1m-x_2m=(x_1-x_2)m \in \lambda(m)$ because $x_1, x_2 \in \lambda$ and λ is a left ideal implies that $x_1-x_2 \in \lambda$. Thus λm is an additive subgroup of M.

Now let $r \in R$ and $xm \in \lambda m$. Then $x \in \lambda$ We have $r(xm) = (rx)m \in \lambda m$ because $r \in R$, $x \in \lambda$ and λ is a left-ideal implies that $rx \in \lambda$. Hence λm is a submodule of M.

Exercises

1. Suppose that R is a ring with unity and that M is a module over R but is not unital. Prove that there exists an $m \neq 0$ in M such that $rm = 0$ for all $r \in R$.
2. If M is an irreducible R-module prove that either M is cyclic or that for every $m \in M$ and $r \in R$, $rm = 0$.
3. If A and B are submodules of M prove that $(A+B)/B$ is isomorphic to $A/(A \cap B)$.
4. Let M, N, Q be three R-modules, and let T be a homomorphism of M into N and S a homomorphism of N into Q. Define $ST : M \to Q$ by $(ST)(m) = [T(m)]$ for any $m \in M$. Prove that ST is an R-homomorphism of M into Q and determine its kernel, $K(ST)$.
5. Let M be an R-module and let $E(M)$ be the set of all R-homomorphisms of M into M. Make appropriate definitions of addition and multiplication of elements of $E(M)$ so that $E(M)$ becomes a ring.

9

Extension Fields and Galois Theory

A field is a commutative ring with unit element in which every non-zero element possesses a multiplicative inverse. If F is a field and $0 \neq a$, $b \in F$, then $a^{-1} b$ is sometimes also written as b/a. There is no ambiguity in writing $a^{-1} b$ in this form because in a field $a^{-1}b = ba^{-1}$. We say that b/a is an element of F obtained on dividing b by a. Thus a field is a commutative ring in which we can divide by any non-zero element. This chapter will be devoted to a study of the theory of finite field extensions and also we shall have some discussion on Galois theory.

§ 1. **Field extensions. Definition.** *Suppose F is a field. Then a field K is said to be an extension of F if F is a subfield of K.*

(Meerut 1981, 84)

In the chapter on vector spaces we have shown that if F is a subfield of a field K, then K can be regarded as a vector space over F under the ordinary field operations in K The dimension of the vector space $K(F)$ will play an important role in this chapter. Throughout this chapter K will denote an extension of F.

Degree of a field extension. Definition.

(Meerut 1978, 81, 84 P ; Vikram 77)

Let K be an extension of the field F. The dimension of K as a vector space over F i.e., the dimension of the vector space $K(F)$ is called the degree of K over F. We shall always denote the degree of K over F by $[K : F]$.

Finite field extension. Definition. **(Kanpur 1971)**

Let K be an extension of the field F. Then K is said to be a finite extension of F if the degree of K over F is finite. Thus K is a finite extension of F if the vector space $K(F)$ is finite dimensional.

In this chapter we shall devote particular attention to finite field extensions because of their importance in the study of the theory of equations. Before proceeding further we want to give some illustrations of field extensions.

Illustrations :

1. If F is any field then F can be regarded as a subfield of F. Therefore F can be thought of as an extension of F. The dimension of the vector space $F(F)$ is one. In fact the unit element 1 of F is a basis of this vector space. Thus the degree of F over F is one *i.e.*, $[F : F]=1$.

2. The field **C** of complex numbers is a finite extension of the field **R** of real numbers. Also we have $[\mathbf{C} : \mathbf{R}]=2$. The set $\{1, i\}$ where $i=\sqrt{(-1)}$ is a basis of the vector space **C**(**R**). If $a, b \in \mathbf{R}$, then

$$a1+bi=0 \Rightarrow a=0,\ b=0.$$

Therefore the set $\{1, i\}$ is linearly independent over **R**.

Also if $a+ib \in \mathbf{C}$, then $a+ib$ is a linear combination of $1, i$ over **R**. Thus the set $\{1, i\}$ generates **C**(**R**). Therefore the set $\{1, i\}$ is a basis for the vector space **C**(**R**). **(Meerut 1980)**

3. Let **Q** be the field of rational numbers. The field $\mathbf{Q}(\sqrt{2})=\{a+b\sqrt{2} : a, b \in \mathbf{Q}\}$ is a finite extension of **Q**. We have $[\mathbf{Q}(\sqrt{2}) : \mathbf{Q}]=2$. In fact the set $\{1, \sqrt{2}\}$ is a basis of $\mathbf{Q}(\sqrt{2})$ regarded as a vector space over the field **Q**

4. The field

$$\mathbf{Q}(\sqrt{2}, \sqrt{3})=\{a+b\sqrt{2}+c\sqrt{3}+d\sqrt{2}\sqrt{3} : a, b, c, d \in \mathbf{Q}\}$$

is a finite extension of **Q**. We have $[\mathbf{Q}(\sqrt{2}, \sqrt{3}) : \mathbf{Q}]=4$. As can be easily seen the set $\{1, \sqrt{2}, \sqrt{3}, \sqrt{2}\sqrt{3}\}$ is a basis of $\mathbf{Q}(\sqrt{2}, \sqrt{3})$ thought of as a vector space over the field **Q**. **(Meerut 1981, 82)**

§ 2. Transitivity of finite extensions.

Theorem 1. *If L is a finite extension of K and if K is a finite extension of F, then L is a finite extension of F. Moreover,*

$$[L : F]=[L : K]\ [K : F].$$

(Banaras 1970; Kurushetra 70; Meerut 79, 80, 82, 83P, 84P; G.N.D.U. Amritsar 82; Vikram 76)

Proof. Let K be a subfield of L and F be a subfield of K *i.e.*, $L \supseteq K \supseteq F$. Let $[L : K]=m$ and $[K : F]=n$.

Suppose $\alpha_1, \alpha_2, \ldots, \alpha_m$ is a basis of L over K and $\beta_1, \beta_2, \ldots, \beta_n$ is a basis of K over F. Then $\alpha_1, \ldots, \alpha_m \in L$ and $\beta_1, \ldots, \beta_n \in K$. Since $K \subseteq L$, therefore $\beta_1, \ldots, \beta_n \in L$. Consequently the mn elements $\alpha_i\beta_j$ where $i=1, \ldots, m$, $j=1, \ldots, n$ are all in L. We shall prove that the set of these mn elements forms a basis of L over F. Then we shall have $[L : F]=mn$ *i.e.*, L will also be a finite extension of F and also we shall have $[L : F]=[L : K]\ [K : F]$.

Thus the theorem shall be proved.

First we shall show that the set $\{\alpha_i\beta_j\}$ generates L over F. Let

γ be any element in L. Since $\{\alpha_1,\ldots,\alpha_m\}$ is a basis of $L(K)$, therefore γ can be expressed as a linear combination of $\alpha_1,\ldots,\alpha_m$ over the elements in K. So we have

$$\gamma=\sum_{i=1}^{m} k_i\alpha_i,\ k_i\in K. \quad \ldots(1)$$

Now $k_i\in K$ and $\{\beta_1,\ldots,\beta_n\}$ is a basis of $K(F)$. Therefore we have

$$k_i=\sum_{j=1}^{n} f_{ij}\,\beta_j,\ f_{ij}\in F. \quad \ldots(2)$$

From (1) and (2), we have

$$\gamma=\sum_{i=1}^{m}\left(\sum_{j=1}^{n} f_{ij}\,\beta_j\right)\alpha_i=\sum_{i=1}^{m}\sum_{j=1}^{n} f_{ij}\,(\alpha_i\beta_j),\ f_{ij}\in F.$$

Thus γ is a linear combination of the elements $\alpha_i\beta_j$ over F. Therefore the set of mn elements $\alpha_i\beta_j$ generates the vector space $L(F)$. Now we shall show that the set $\{\alpha_i\beta_j\}$ is linearly independent over F. We have

$$\sum_{i=1}^{m}\sum_{j=1}^{n} f_{ij}\,(\alpha_i\beta_j)=0,\ f_{ij}\in F \Rightarrow \sum_{i=1}^{m}\left(\sum_{j=1} f_{ij}\,\beta_j\right)\alpha_i=0$$

$\Rightarrow \sum_{j=1} f_{ij}\,\beta_j=0$ for $i=1,\ldots,m$, since $\{\alpha_1,\ldots,\alpha_m\}$ is a basis of $L(K)$ and each $f_{ij}\,\beta_j\in K$

$\Rightarrow f_{ij}=0$ for $i=1,\ldots,m,\ j=1,\ldots,n$ since $\{\beta_1,\ldots,\beta_n\}$ is a basis of $K(F)$ and each $f_{ij}\in F$

$\Rightarrow$ the set $\{\alpha_i\beta_j\}$ is linearly independent over F.

Hence the set $\{\alpha_i\beta_j\}$ is a basis of L over F. This proves the theorem.

Theorem 2. *If L is a finite extension of F and if K is a subfield of L which contains F, then $[K:F]\mid[L:F]$ i.e., $[K:F]$ is a divisor of $[L:F]$.*

Proof Let L, K, F be three fields in the relation $L\supseteq K\supseteq F$. Suppose further that $[L:F]$ is finite and is equal to n. Let $\{\alpha_1,\ldots,\alpha_n\}$ be a basis of L over F. Then $\{\alpha_1,\ldots,\alpha_n\}$ generates L over F. Since $K\supseteq F$, therefore any linear combination of $\alpha_1,\ldots,\alpha_n$ over F will also be a linear combination of $\alpha_1,\ldots,\alpha_n$ over K. Therefore the set $\{\alpha_1,\ldots,\alpha_n\}$ also generates L over K though it may not be linearly independent over K. Since $L(K)$ is generated by a finite set, therefore it is a finite dimensional vector space and so $[L : K]$ is finite. Further $K(F)$ is a subspace of $L(F)$. Since $[L : F]$ is finite therefore $[K : F]$ is finite. Recall that each subspace of a finite dimensional vector space is also finite dimensional. Now by theorem (1), we have

$[L : F]=[L : K]\ [K : F] \Rightarrow [K : F]$ is a divisor of $[L : F]$.

Note. From theorem (2) we conclude that if $[L : F]$ is a prime number, then there can be no field K properly contained between L and F. In other words if $[L : F]$ is prime and K is any subfield of L containing F, then either we have $K=L$ or we have $K=F$.

§ 3 **Field adjunctions.** Suppose K is an extension of a field F. Let $a \in K$. Suppose C is the collection of all subfields of K containing both F and a. C is not empty because at least K itself is in C. Now the intersection of an arbitrary collection of subfields of K is also a subfield of K. Let $F(a)$ denote the intersection of all those subfields of K which are members of C. Then $F(a)$ is a subfield of K. Obviously $F(a)$ contains both F and a because each member of C contains both F and a. Thus $F(a)$ is a member of C. Further if E is any subfield of K containing both F and a, then $F(a)$ will be contained in E. The reason is that $F(a)$ is the intersection of the members of C and E is a member of C. Thus $F(a)$ is a subfield of K containing both F and a and itself is contained in any subfield of K containing both F and a. *Therefore $F(a)$ is the smallest subfield of K containing both F and a.* We call $F(a)$ the subfield obtained by *adjoining a to F*. Here a has been adjoined to F and the process is called *field adjunction.*

Constructive description of $F(a)$. Suppose K is an extension of a field F. Let $a \in K$. Let $U=\left\{\dfrac{k_0a^n+k_1a^{n-1}+\ldots+k_n}{l_0a^m+l_1a^{m-1}+\ldots+l_m}\right.$: the k_i and l_i are elements in F, $l_0a^m+l_1a^{m-1}+\ldots+l_m$ is not equal to the zero element of K and n, m are any non-negative integers $\left.\right\}$.

Obviously U is a subfield of K. It can be easily seen that

(i) $\alpha, \beta \in U \Rightarrow \alpha-\beta \in U$

(ii) $\alpha \in U, 0 \neq \beta \in U \Rightarrow \alpha/\beta \in U.$

Thus U is a subfield of K. We claim that $U=F(a)$.

Obviously U contains both F and a. Therefore U is a subfield of K containing both F and a. This implies that $U \supseteq F(a)$.

Further any subfield of K which contains both F and a, by virtue of closure under addition and multiplication, must contain all the elements $k_0a^n+k_1a^{n-1}+...+k_n$ where each $k_i \in F$. Since $F(a)$ is a subfield of K containing both F and a, therefore $F(a)$ must contain all such elements. Being a subfield of K, $F(a)$ must also contain all quotients of such elements. Therefore $F(a)$ must contain U *i e.*, $F(a) \supseteq U$.

Now $U \supseteq F(a)$ and $F(a) \supseteq U \Rightarrow U=F(a)$.

Simple field extension. Definition. *The extension K of a field F is called a simple extension of F if $K=F(a)$ for some a in K.*

(Meerut 1981, 82, 83)

Let K be an extension of a field F. Let $a, b \in K$. Let $T=F(a)$. Since $F(a)$ is a subfied of K, therefore K is also an extension of $F(a)$. Let W be the subfield of K obtained by adjoining b to $F(a)$. Then $W=(F(a))(b)$. We shall write $(F(a))(b)$ as $F(a, b)$. Similarly we can describe $F(b, a)$. We have

$F(a, b)=(F(a))(b)$

$=$the smallest subfield of K conatining both $F(a)$ and b

$=$the smallest subfield of K containing F, a and b because any subfield of K which contains both F and a must contain $F(a)$.

Similarly $F(b, a)=$the smallest subfield of K containing F, a and b.

Since the subfield of K generated by F, a and b is unique, therefore $F(a, b)=F(b, a)$.

Thus $F(a, b)$ is the subfield of K obtained by adjoining both a and b to F.

Similarly if $a_1, a_2, \cdots, a_n \in K$, then $F(a_1, a_2, \cdots, a_n)$ will be described as the subfield of K generated by F, $a_1, \cdots, a_n$. In other words $F(a_1, a_2, \cdots, a_n)$ will be the smallest subfield of K containing F as well as $a_1, a_2, \cdots, a_n$.

§ 4. Algebraic field extensions.

Let $q(x) \in F[x]$, the ring of polynomials in x over F. Let

$$q(x)=\alpha_0x^m+\alpha_1x^{m-1}+..+\alpha_m.$$

Suppose $b \in K$ where K is any extension of F. Then by $q(b)$ we shall mean the element $\alpha_0b^m+\alpha_1b^{m-1}+...+\alpha_m$ in K. Sometimes $q(b)$ is also called the value of $q(x)$ obtained by substituting b for x. The element b is said to *satisfy* $q(x)$ if $q(b)=0$. Also then we say that b is a *root of* $q(x)$.

Algebraic element. Definition. *Let K be an extension of a field F. An element $a \in K$ is said to be algebraic over F if there is* *a non-zero polynomial* $p(x) \in F[x]$ *for which* $p(a)=0$. **(Kanpur 1980)**

In other words $a \in K$ is said to be algebraic over F if there exist elements $\beta_0, \beta_1, \ldots, \beta_n$ in F, not all 0, such that

$$\beta_0 a^n + \beta_1 a^{n-1} + \ldots + \beta_n = 0.$$

Transcendental element. Definition. *Let K be an extension of a field F. An element* $a \in K$ *is said to be transcendental over F if it is not algebraic over F.*

Definition. *A complex number is said to be an algebraic number if it is algebraic over the field of rational numbers.*

A complex number which is not algebraic is called *transcendental* The number e is transcendental.

Minimal polynomial of an algebraic element. Definition. *Let K be an extension of a field F. Let* $a \in K$ *be algebraic over F. Suppose* $p(x)$ *is a polynomial over F of lowest positive degree satisfied by a. Then* $p(x)$ *is called a minimal polynomial for a over F.*

Let us impose the restriction on minimal polynomial for a over F that it should be monic *i e.*, that in it the coefficient of highest power of x should be 1. Then we can speak of as the minimal polynomial for a over F because it will be unique.

Theorem 3. *Let* $a \in K$ *be algebraic over F. Then any two minimal monic polynomials for a over F are equal.*

Proof Let $x^n + \alpha_1 x^{n-1} + \ldots + \alpha_n$ and $x^n + \beta_1 x^{n-1} + \ldots + \beta_n$ be two minimal monic polynomials for a over F. Then

$a^n + \alpha_1 a^{n-1} + \ldots + \alpha_n = 0 = a^n + \beta_1 a^{n-1} + \ldots + \beta_n$

$\Rightarrow a^n + \alpha_1 a^{n-1} + \ldots + \alpha_n = a^n + \beta_1 a^{n-1} + \ldots + \beta_n$

$\Rightarrow (\alpha_1 - \beta_1) a^{n-1} + (\alpha_2 - \beta_2) a^{n-2} + \ldots + (\alpha_n - \beta_n) = 0$

$\Rightarrow$ a satisfies the polynomial $q(x) = (\alpha_1 - \beta_1) x^{n-1} + \ldots + (\alpha_n - \beta_n)$ belonging to $F[x]$

$\Rightarrow$ $q(x)$ must be the zero polynomial because minimal polynomial for a over F is of degree n while $q(x)$, if it is not the zero polynomial, is of degree less than n

$\Rightarrow \alpha_1 - \beta_1 = 0, \ldots, \alpha_n - \beta_n = 0 \Rightarrow \alpha_1 = \beta_1, \ldots, \alpha_n = \beta_n$

$\Rightarrow x^n + \alpha_1 x^{n-1} + \ldots + \alpha_n = x^n + \beta_1 x^{n-1} + \ldots + \beta_n.$

This completes the proof of the theorem.

Irreducibility of minimal polynomial.

Theorem 4. *Let* $a \in K$ *be algebraic over F and let* $p(x)$ *be a minimal polynomial for a over F. Then* $p(x)$ *is irreducible over F.*

Proof. Suppose $p(x)$ is a polynomial in $F[x]$ of smallest positive degree such that $p(a)=0$. Suppose $p(x)$ is not irreducible over F. Then $p(x)$ can be resolved into non-trivial factors. Let $p(x)=f(x)\,g(x)$ where $f(x)$ and $g(x)$ are polynomials of positive degree in $F[x]$ and each of them is of degree less than that of $p(x)$. We have

$$\begin{aligned} p(a)=f(a)\,g(a) &\Rightarrow 0=f(a)\,g(a) \qquad [\because \ a \text{ satisfies } p(x)] \\ &\Rightarrow f(a)=0 \text{ or } g(a)=0 \\ &\Rightarrow a \text{ satisfies } f(x) \text{ or } g(x) \\ &\Rightarrow p(x) \text{ is not a minimal polynomial for } a \text{ over } F \text{ because} \\ &\qquad \deg f(x) < \deg p(x) \text{ and } \deg g(x) < \deg p(x). \end{aligned}$$

Since $p(x)$ is a minimal polynomial for a over F, therefore our assumption that $p(x)$ is not irreducible over F is wrong. Hence $p(x)$ must be irreducible over F.

Degree of an algebraic element. Definition. *Let K be an extension of the field F. The element $a \in K$ is said to be algebraic of degree n over F if it satisfies a non-zero polynomial over F of degree n but no non-zero polynomial of lower degree.*

Thus $a \in K$ is algebraic of degree n over F if the minimal polynomial for a over F is of degree n.

Algebraic extension. Definition. *The extension K of F is called an algebraic extension of F if every element in K is algebraic over F.*

If there exists $a \in K$ such that a is not algebraic over F, then K is called a *transcendental extension of F*.

The field $\mathbf{C}$ of complex numbers is an algebraic extension of $\mathbf{R}$, the field of real numbers.

The field $\mathbf{R}$ of real numbers is not an algebraic extension of the field $\mathbf{Q}$ of rational numbers. In fact π is an element of $\mathbf{R}$ which is not algebraic over $\mathbf{Q}$. It was proved by Hermite in 18/1 that there exists no non-zero polynomial with rational coefficients satisfied by π.

If F is any field, then F is an algebraic extension of F. If $a \in F$, then a satisfies non-zero polynomial $1x-a$ in $F[x]$.

Theorem 5. *Let K be an extension of a field F and let $a \in K$ be algebraic of degree n over F. Then*

$$F(a)=\{\beta_0+\beta_1 a+\beta_2 a^2+\ldots+\beta_{n-1} a^{n-1} : \beta_0, \beta_1, \ldots, \beta_{n-1} \in F\}.$$

Also the expression for each element of $F(a)$ in the form $\beta_0+\beta_1 a+\ldots+\beta_{n-1} a^{n-1}$ is unique.

Proof. If $n=1$, then $a \in F$. Therefore in this case $F(a)=F$ and the result of the theorem is obviously true.

So let us take $n > 1$. Since $a \in K$ is algebraic of degree n over F, therefore the minimal polynomial for a over F is of degree n. Let $p(x) = x^n + \alpha_1 x^{n-1} + \dots + \alpha_n$ be the minimal polynomial for a over F. Then

$$a^n + \alpha_1 a^{n-1} + \dots + \alpha_n = 0$$

$$\Rightarrow a^n = -(\alpha_1 a^{n-1} + \dots + \alpha_n) \qquad \dots(1)$$

$$\Rightarrow a^{n+1} = -(\alpha_1 a^n + \alpha_2 a^{n-1} + \dots + \alpha_n a)$$

[on multiplying both sides by a]

$$\Rightarrow a^{n+1} = -[-\alpha_1 (\alpha_1 a^{n-1} + \dots + \alpha_n) + \alpha_2 a^{n-1} + \dots + \alpha_n a]$$

[putting the value of a^n from (1)]

$\Rightarrow a^{n+1}$ is a linear combination of the elements $1, a, \dots, a^{n-1}$ over F.

Continuing the above process we can show that a^{n+k}, for $k > 0$, is a linear combiration over F of $1, a, \dots, a^{n-1}$.

Now let $T = \{\beta_0 + \beta_1 a + \dots + \beta_{n-1} a^{n-1} : \beta_0, \beta_1, \dots, \beta_{n-1} \in F\}$.

We shall show that $T = F(a)$. For this we shall first prove that T is a subfield of K.

Let $u = \beta_0 + \beta_1 a + \dots + \beta_{n-1} a^{n-1}$, $v = \gamma_0 + \gamma_1 a + \dots + \gamma_{n-1} a^{n-1}$ be any two elements in T. Then

$$u - v = (\beta_0 - \gamma_0) + (\beta_1 - \gamma_1) a + \dots + (\beta_{n-1} - \gamma_{n-1}) a^{n-1} \in T.$$

Now let $0 \neq u = \beta_0 + \beta_1 a + \dots + \beta_{n-1} a^{n-1}$ be in T.

Let $q(x) = \beta_0 + \beta_1 x + \dots + \beta_{n-1} x^{n-1} \in F[x]$. Then $q(a) = u \neq 0$.

We claim that $q(x)$ is not a divisor of $p(x)$. Because if $q(x)$ is a divisor of $p(x)$, then we must have $p(x) = (a_0 x + a_1)\, q(x)$ where $a_0 x + a_1$ is a non-zero polynomial in $F[x]$.

Putting $x = a$ in this relation, we get

$$p(a) = (a_0 a + a_1)\, q(a)$$

$$\Rightarrow 0 = (a_0 a + a_1)\, q(a) \qquad [\because p(a) = 0]$$

$$\Rightarrow a a_0 + a_1 = 0 \qquad [\because q(a) \neq 0]$$

$\Rightarrow a$ satisfies a polynomial $a_0 x + a_1$ of degree 1 over F because $\deg p(x) = n$ which is > 1.

Therefore $q(x)$ is not a divisor of $p(x)$. Since $p(x)$ is irreducible over F, therefore $q(x)$ and $p(x)$ must be relatively prime. Consequently we can find polynomials $s(x)$ and $t(x)$ in $F[x]$ such that $p(x)\, s(x) + q(x)\, t(x) = 1$. Putting $x = a$ in this relation, we get

$$p(a)\, s(a) + q(a)\, t(a) = 1$$

$$\Rightarrow q(a)\, t(a) = 1 \qquad [\because p(a) = 0]$$

$$\Rightarrow u t(a) = 1 \qquad [\because q(a) = u]$$

$\Rightarrow t(a)$ is the inverse of u.

Now in $t(a)$ all powers of a higher than $n-1$ can be replaced by linear combinations of $1, a, \ldots, a^{n-1}$ over F.

Therefore $t(a) \in T$. Thus $t(a)=u^{-1} \in T$.

Now in the product $u^{-1}v$ all powers of a higher than $n-1$ can be replaced by linear combinations of $1, a, \ldots, a^{n-1}$ over F.

Therefore $0 \neq u,\ v \in T \Rightarrow u^{-1}v \in T$. Thus T is a subfield of K.

Now from the definition of T it is obvious that both F and a are in T. Also by virtue of closure under addition and multiplication any subfield of K which contains both F and a must contain T. Thus T is the smallest subfield of K containing both F and a. Hence $$T=F(a).$$

Now let $u \in T$. Further let $u=\beta_0+\beta_1 a+\ldots+\beta_{n-1} a^{n-1}$ and also $u=\gamma_0+\gamma_1 a+\ldots+\gamma_{n-1} a^{n-1}$. Then

$$\beta_0+\beta_1 a+\ldots+\beta_{n-1} a^{n-1}=\gamma_0+\gamma_1 a+\ldots+\gamma_{n-1} a^{n-1}$$

$\Rightarrow (\beta_0-\gamma_0)+(\beta_1-\gamma_1)\, a+\ldots+(\beta_{n-1}-\gamma_{n-1})\, a^{n-1}=0$

$\Rightarrow$ a satisfies the polynomial

$$h(x)=(\beta_0-\gamma_0)+(\beta_1-\gamma_1)\, x+\ldots+(\beta_{n-1}-\gamma_{n-1})\, x^{n-1}$$

belonging to $F[x]$

$\Rightarrow$ $h(x)$ must be the zero polynomial because otherwise a will not be of degree n over F

[Note that if $h(x) \neq 0$, then deg $h(x) < n$]

$\Rightarrow \beta_0-\gamma_0=0, \beta_1-\gamma_1=0, \ldots, \beta_{n-1}-\gamma_{n-1}=0$

$\Rightarrow \quad \beta_0=\gamma_0, \quad \beta_1=\gamma_1, \ldots, \beta_{n-1}=\gamma_{n-1}$

$\Rightarrow$ the expression for u in the form $\beta_0+\beta_1 a+\ldots+\beta_{n-1} a^{n-1}$ is unique.

Theorem 6. *Let K be an extension of a field F and let $a \in K$ be algebraic over F. Suppose a satisfies an irreducible polynomial $p(x)$ in $F[x]$. Then $p(x)$ must be a minimal polynomial for a over F.*

Proof. Let $M=\{f(x) \in F[x] : f(a)=0\}$. We claim that M is an ideal of $F[x]$. The proof is as follows :

Let $f(x), g(x) \in M$. Then $f(a)=0, g(a)=0$.

Let $s(x)=f(x)-g(x)$. Then $s(a)=f(a)-g(a)=0-0=0$.

$\therefore \quad s(x) \in M$.

Also let $f(x) \in M$ and $h(x) \in F[x]$. Then $f(a)=0$.

Let $t(x)=f(x) . h(x)$. Then $t(a)=f(a) . h(a)=0 . h(a)=0$.

$\therefore \quad t(x) \in M$.

Hence M is an ideal of $F[x]$. Obviously $M \neq F[x]$. Because if $f(x)=1 \in F[x]$, then $f(a)=1 \neq 0$.

Then $f(x)=1$ is not in M. Therefore $M \neq F[x]$.

Now $p(x)$ is an irreducible polynomial in $F[x]$. Therefore the ideal $N=(p(x))$ of $F[x]$ generated by $p(x)$ is a maximal ideal because $F[x]$ is a Euclidean ring. We have $p(x) \in M$ because it is given that $p(a)=0$. Now if $n(x)=m(x)\,p(x)$ is any element of $(p(x))$, then $n(a)=m(a)\,p(a)=0$. This implies that $n(x)$ is in M.

Thus $N \subseteq M$. Therefore M is an ideal of $F[x]$ contained between N and $F[x]$ *i.e.*, $N \subseteq M \subseteq F[x]$.

Since N is a maximal ideal and $M \neq F[x]$, therefore we must have $N=M \Rightarrow M=(p(x))$.

Now suppose that $p(x)$ is not a minimal polynomial for a over F. Let $q(x)$ be a polynomial in $F[x]$ of degree lower than that of $p(x)$ and satisfied by a.

Since $q(a)=0$, therefore $q(x) \in M$.

$\therefore$ $q(x)=p(x)\,r(x)$ for some $r(x) \in F[x]$.

But this result is absurd because $\deg q(x) < \deg p(x)$.

Hence $p(x)$ must be a minimal polynomial for a over F.

Theorem 7. *Let K be an extension of a field F. Then the element $a \in K$ is algebraic over F if and only if $F(a)$ is a finite extension of F.*

(I.A.S. 1971 ; Kanpur 80 ; Meerut 77, 78, 79, 81, 82, 83, 84; Banaras 72 ; Calicut 75)

Proof. Suppose $F(a)$ is a finite extension of F. Then to prove that a is algebraic over F. Let $[F(a):F]=m$. Since $F(a)$ is a field and $a \in F(a)$, therefore the $m+1$ elements

$$1,\ a,\ a^2, \ldots,\ a^{m-1},\ a^m$$

are all in $F(a)$. Since the dimension of the vector space $F(a)$ over F is m, therefore these $m+1$ elements of $F(a)$ are linearly dependent over F. So there exist elements $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_m \in F$, not all 0, such that

$$\alpha_0 1+\alpha_1 a+\alpha_2 a^2+\ldots+\alpha_m a^m=0$$

$\Rightarrow$ a satisfies a non-zero polynomial $f(x)=\alpha_0+\alpha_1 x+\alpha_2 x^2+\ldots+\alpha_m x^m \in F[x]$

$\Rightarrow$ a is algebraic over F.

This proves the 'if' part of the theorem.

Now we shall prove the 'only 'if' part of the theorem. It is given that $a \in K$ is algebraic over F and we are to prove that $F(a)$ is a finite extension of F. Let $p(x)$ be a polynomial over F of lowest positive degree satisfied by a. Let degree $p(x)=n$. Then a is algebraic of degree n over F. Therefore

$$F(a)=\{\beta_0+\beta_1 a+\beta_2 a^2+\ldots+\beta_{n-1} a^{n-1} : \beta_0, \beta_1, \ldots, \beta_{n-1} \in F\}.$$

From this we see that $F(a)$ is a vector space over F spanned by the elements $1, a, a^2, \ldots, a^{n-1}$. These elements of $F(a)$ are also linearly independent over F. Because

$$\gamma_0 1 + \gamma_1 a + \gamma_2 a^2 + \ldots + \gamma_{n-1} a^{n-1} = 0, \text{ with } \gamma_i \in F$$

$\Rightarrow$ a satisfies a polynomial $q(x) = \gamma_0 + \gamma_1 x + \ldots + \gamma_{n-1} x^{n-1}$ belonging to $F[x]$

$\Rightarrow$ $q(x)$ must be the zero polynomial otherwise $p(x)$ will not be a polynomial of lowest positive degree satisfied by a

[Note that deg $p(x) = n$ while deg $q(x) < n$ if $q(x) \neq 0$]

$\Rightarrow$ $\gamma_0 = 0, \gamma_1 = 0, \gamma_2 = 0, \ldots, \gamma_{n-1} = 0$

$\Rightarrow$ $1, a, a^2, \ldots, a^{n-1} \in F(a)$ are linearly independent over F.

Thus the n elements $1, a, a^2, \ldots, a^{n-1}$ constitute a basis for $F(a)$ over F. Hence $[F(a) : F] = n$. Therefore $F(a)$ is a finite extension of F.

Remark. If $[F(a) : F] = m$, a is algebraic over F. Let the degree of a over F be n. Then we have $[F(a) : F] = n$. Therefore $m = n$. Hence if $[F(a) : F] = m$, then a is algebraic of degree m over F.

Theorem 8. *Let K be an extension of a field F and let $a \in K$ be algebraic of degree n over F. Then $[F(a) : F] = n$.*

Proof. This theorem is nothing but the 'only if' part of the previous theorem.

Theorem 9. *Every finite extension K of a field F is algebraic.*

(Banaras 1970 ; Gujrat 76 ; Meerut 79)

Proof Suppose K is a finite extension of F. Then K will be an algebraic extension of F if every element a in K is algebraic over F. Let $[K : F] = m$. Since K is a field and $a \in K$, therefore the $m+1$ elements $1, a, a^2, \ldots, a^{m-1}, a^m$ are all in K. Since the dimension of the vector space $K(F)$ is m, therefore these $m+1$ elements of K are linearly dependent over F. So there exist elements $\alpha_0, \alpha_1, \ldots, \alpha_m \in F$, not all zero, such that

$$\alpha_0 1 + \alpha_1 a + \alpha_2 a^2 + \ldots + \alpha_m a^m = 0$$

$\Rightarrow$ a satisfies a non-zero polynomial

$$p(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_m x^m \in F[x] \text{ of degree at most } m$$

$\Rightarrow$ a is algebraic over F.

Hence K is an algebraic extension of F.

Remark. If $[K ; F] = m$, then each element a in K is algebraic over F and the degree of a over F will be $\leqslant m$.

Theorem 10. *Let K be an extension of a field F and let $a_1, a_2, \ldots, a_n$ be n elements in K algebraic over F.*

Then $F(a_1, a_2, \ldots, a_n)$ *is a finite extension of* F *and consequently an algebraic extension of* F. **(Meerut 1976)**

Proof. We have

$$F \subseteq F(a_1) \subseteq F(a_1, a_2) \subseteq \cdots \subseteq F(a_1, a_2, \ldots, a_n) \subseteq K.$$

Since a_k is algebraic over F, therefore it is also algebraic over $F(a_1, a_2, \ldots, a_{k-1})$ which is a superfield of F. Note that any non-zero polynomial over F is also a non-zero polynomial over

$$F(a_1, a_2, \ldots, a_{k-1}).$$

Now a_k is algebraic over $F(a_1, a_2, \ldots, a_{k-1})$

$\Rightarrow$ $(F(a_1, a_2, \ldots, a_{k-1}))(a_k)$ is a finite extension of

$$F(a_1, a_2, \ldots, a_{k-1})$$

$\Rightarrow$ $F(a_1, a_2, \ldots, a_{k-1}, a_k)$ is a finite extension of

$$F(a_1, a_2, \ldots, a_{k-1})$$

$\Rightarrow$ $[F(a_1, a_2, \ldots, a_k) : F(a_1, a_2, \ldots, a_{k-1})]$ is finite, say λ_k.

Now $[F(a_1, a_2, \ldots, a_n) : F]$

$=[F(a_1, a_2, \ldots, a_n) : F(a_1, a_2, \ldots a_{n-1})]$

$$[F(a_1, a_2, \ldots, a_{n-1}) : F(a_1, a_2, \ldots, a_{n-2})] \ldots [F(a_1) : F]$$

$=\lambda_n \lambda_{n-1} \ldots \lambda_2 \lambda_1 =$ finite since each λ is finite.

Hence $F(a_1, a_2, \ldots, a_n)$ is a finite extension of F. Consequently $F(a_1, a_2, \ldots, a_n)$ is an algebraic extension of F

Theorem 11. *Let K be an extension of a field F. Then the elements in K which are algebraic over F form a subfield of K. In other words if a, b in K are algebraic over F, then* $a \pm b$, ab, *and* a/b *(if* $b \neq 0$*) are all algebraic over F.* **(Banaras 1972; Meerut 73, 76)**

Proof. Suppose $a, b \in K$ are algebraic over F. Since b is algebraic over F, therefore it is also algebraic over $F(a)$ which is a superfield of F. Note that any non-zero polynomial over F is also a non-zero polynomial over $F(a)$.

Now b is algebraic over $F(a) \Rightarrow (F(a))(b)$ *i.e.*, $F(a, b)$ is a finite extension of $F(a)$. Therefore $[F(a, b) : F(a)] =$ finite.

Also $[F(a) : F]$ is finite because a is algebraic over F.

Now $[F(a, b) : F] = [F(a, b) : F(a)] [F(a) : F] =$ finite.

$\therefore$ $F(a, b)$ is a finite extension of F. Consquently $F(a, b)$ is an algebraic extension of F. Now $F(a, b)$ is a field and $a, b \in F(a, b)$. Therefore $a \pm b$, ab and a/b (if $b \neq 0$) are all in $F(a, b)$. So $a \pm b$, ab, and a/b are all algebraic over F.

Hence the elements algebraic over F form a subfield of K.

Theorem 12. *If a and b in K are algebraic over F of degrees m and n respectively, then* $a \pm b$, ab *and* a/b *(if* $b \neq 0$*) are algebraic over F of degrees at most mn.* **(Banaras 1969)**

Proof. Since a is algebraic of degree m over F, therefore $F(a)$ is of degree m over F *i e.*, $[F(a) : F] = m$. Again b is algebraic of degree n over F. Therefore b is algebraic of degree at most n over $F(a)$ which is a superfield of F. This implies that the sub-field $(F(a))(b)$ *i.e.*, $F(a, b)$ of K is of degree at most n over $F(a)$ *i.e.*,

$$[F(a, b) : F(a)] \leqslant n.$$

Now $[F(a, b) : F] = [F(a, b) : F(a)]\,[F(a) : F] \leqslant mn$.

Since $[F(a, b) : F]$ is finite, therefore $F(a, b)$ is a finite extension of F and so it is an algebraic extension of F. Each element in $F(a, b)$ will be algebraic of degree $\leqslant mn$ over F. Since $F(a, b)$ is a field, therefore $a, b \in F(a, b) \Rightarrow a \pm b, ab, a/b$ (if $b \neq 0$) are all in $F(a, b)$. Hence $a \pm b, ab, a/b$, (if $b \neq 0$) are algebraic of degree at most mn over F.

Theorem 13. Transitivity of Algebraic extension. *If L is an algebraic extension of K and K is an algebraic extension of F, then L is an algebraic extension of F.* **(I.A.S. 1970 ; Meerut 78, 84)**

Proof. Let a be any arbitrary element in L. If we prove that a is algebraic over F, then L will be an algebraic extension of F.

Since $a \in L$ and L is an algebraic extension of K, therefore a satisfies some polynomial $x^n + \alpha_1 x^{n-1} + \alpha_2 x^{n-2} + \ldots + \alpha_n$ where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are in K. Now K is an algebraic extension of F. Therefore $\alpha_1, \alpha_2, \ldots, \alpha_n$ are algebraic over F. So

$$M = F(\alpha_1, \alpha_2, \ldots, \alpha_n)$$

is a finite extension of F. Now a satisfies the polynomial $x^n + \alpha_1 x^{n-1} + \ldots + \alpha_n$ whose coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$ are in $M = F(\alpha_1, \alpha_2, \ldots, \alpha_n)$. Therefore a is algebraic over M. Consequently $M(a)$ is a finite extension of M.

Now $M(a)$ is a finite extension of M and M is a finite extension of F. Therefore $M(a)$ is a finite extension of F. So a is algebraic over F. This completes the proof of the theorem.

Solved Examples

Ex. 1. *Let K be an extension of a field F. Prove that the mapping $\phi : F[x] \to F(a)$ defined by $h(x)\,\psi = h(a)$ is a homomorphism. Here by $h(x)\,\psi$ we mean the image of $h(x)$ under the mapping ψ.*

Solution. Let $h(x), g(x) \in F[x]$. Then $h(x)\,\psi = h(a)$ and $g(x)\,\psi = g(a)$.

Let $s(x) = h(x) + g(x)$ and $t(x) = h(x)\,g(x)$.

Then $s(a) = h(a) + g(a)$ and $t(a) = h(a)\,g(a)$.

We have $[h(x) + g(x)]\,\psi = s(x)\,\psi = s(a)$

$$= h(a) + g(a) = h(x)\,\psi + g(x)\,\psi.$$

Also $[h(x)\, g(x)]\, \psi = t(x)\, \psi = t(a)$

$$= h(a)\, g(a) = [h(x)\, \psi]\, [g(x)\, \psi].$$

Hence ψ is a homomorphism from $F[x]$ into $F(a)$.

Ex. 2. *Let F be a field and let $F[x]$ be the ring of polynomials in x over F. Let $g(x)$, of degree n, be in $F[x]$. and let $V = (g(x))$ be the ideal generated by $g(x)$ in $F[x]$. Prove that $F[x]/V$ is an n-dimensional vector space over F.*

Solution. We have $V = \{ f(x)\, g(x) : f(x) \in F[x] \}$.

Also $\quad F[x]/V = \{ V + f(x) : f(x) \in F[x] \}$.

We define addition in $F[x]/V$ as follows :

Let $\quad V + f_1(x),\ V + f_2(x) \in F[x]/V$. Then we define

$$\{V + f_1(x)\} + \{V + f_2(x)\} = V + f_1(x) + f_2(x).$$

Also we define scalar multiplication in $F[x]/V$ over F as follows :

Let $a \in F$ and $V + f(x) \in F[x]/V$. Then we define

$$a\, [V + f(x)] = V + a f(x).$$

Obviously $F[x]/V$ is an abelian group with respect to addition defined on it. The residue class V is the zero vector.

Further let $a, b \in F$ and $f_1(x), f_2(x) \in F[x]$. Then

(i) $(a + b)\, [V + f_1(x)] = V + (a + b)\, f_1(x) = V + a f_1(x) + b f_1(x)$

$$= [V + a f_1(x)] + [V + b f_1(x)] = a\, [V + b f_1(x)] + b\, [V + f_1(x)].$$

(ii) $a\, [\{V + f_1(x)\} + \{V + f_2(x)\}] = a\, [V + f_1(x) + f_2(x)]$

$$= V + a\, \{f_1(x) + f_2(x)\} = V + a f_1(x) + a f_2(x)$$

$$= [V + a f_1(x)] + [V + a f_2(x)] = a\, [V + f_1(x)] + a\, [V + f_2(x)].$$

(iii) $a\, [b\, \{V + f_1(x)\}] = a\, [V + b f_1(x)] = V + (ab)\, f_1(x)$

$$= (ab)\, [V + f_1(x)].$$

(iv) $1\, [V + f_1(x)] = V + 1\, f_1(x) = V + f_1(x)$.

Hence $F[x]/V$ is a vector space over F.

Now if $g(x)$ is of degree n, then to show that $F[x]/V$ is of dimension n over F. We claim that $V + 1,\ V + x,\ V + x^2, \ldots,\ V + x^{n-1}$ constitute a basis of $F[x]/V$ over F.

First we shall show that these n elements of $F[x]/V$ are linearly independent over F. Note that V is the zero vector. Also we recall that $V + f(x) = V \Leftrightarrow f(x) \in V$. Now we have

$$a_0(V + 1) + a_1(V + x) + a_2(V + x^2) + \ldots + a_{n-1}(V + x^{n-1}) = V,$$

$$a_i \in F$$

$$\Rightarrow (V + a_0) + (V + a_1 x) + (V + a_2 x^2) + \ldots + (V + a_{n-1} x^{n-1}) = V$$

$$\Rightarrow V + a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1} x^{n-1} = V$$

$$\Rightarrow a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1} x^{n-1} \in V$$

$\Rightarrow a_0+a_1x+a_2x^2+\ldots+a_{n-1}x^{n-1}=f(x)\,g(x)$

for some $f(x)\in F[x]$

$\Rightarrow f(x)=0$ [$\because$ if $f(x)\neq 0$, then $\deg f(x)\,g(x) \geq \deg g(x)=n$ and so we cannot have $f(x)\,g(x)=a_0+a_1x+\ldots+a_{n-1}x^{n-1}$]

$\Rightarrow a_0+a_1x+a_2x^2+\ldots+a_{n-1}x^{n-1}=0$

$\Rightarrow a_0=0,\ a_1=0,\ldots,\ a_{n-1}=0.$

$\therefore$ $V+1,\ V+x,\ V+x^2,\ldots,\ V+x^{n-1}$ are linearly independent over F.

Now we shall show that $V+1, V+x,\ldots, V+x^{n-1}$ generate $F[x]/V$ over F. Let $V+f(x)$ be any element in $F[x]/V$. Then $f(x)\in F[x]$. By division algorithm there exist $q(x)$, $r(x)\in F[x]$ such that $f(x)=q(x)\,g(x)+r(x)$ where either

$$r(x)=0 \quad \text{or} \quad \deg r(x) < \deg g(x).$$

Now $V+f(x)=V+q(x)\,g(x)+r(x)$

$=[V+q(x)\,g(x)]+[V+r(x)]$

$=V+[V+r(x)]$ [$\because$ $q(x)\,g(x)\in V$]

$=V+r(x)$ [$\because$ V is zero veotor]

$=V+a_0+a_1x+a_2x^2+\ldots+a_{n-1}x^{n-1}$, where $a_0, a_1,\ldots, a_{n-1}\in F$

[$\because$ $r(x)=0$ or $\deg r(x) < n$ *i.e.*, $\deg g(x)$]

$=a_0\,(V+1)+a_1\,(V+x)+\ldots+a_{n-1}\,(V+x^{n-1}).$

Hence $V+1$, $V+x,\ldots, V+x^{n-1}$ form a basis of $F[x]/V$ over F. Therefore dim $F[x]/V$ over $F=n$

Ex 3 (a) *Let R be the field of real numbers and Q the field of rational numbers. In R, $\sqrt{2}$ and $\sqrt{3}$ are both algebraic over Q. Exhibit a polynomial of degree 4 over Q satisfied by $\sqrt{2}+\sqrt{3}$.*

(b) *Show that $Q(\sqrt{2}, \sqrt{3})=Q(\sqrt{2}+\sqrt{3})$.* **(Meerut 1980)**

(c) *What is the degree of $\sqrt{2}+\sqrt{3}$ over Q? Prove your answer.*

Solution (a) The element $\sqrt{2}\in R$ satisfies the polynomial x^2-2 over Q. Also x^2-2 is an irreducible polynomial over Q. Therefore the degree of $\sqrt{2}$ over $Q=$ the degree of $x^2-2=2$. Also $[Q(\sqrt{2}):Q]=$ degree of $\sqrt{2}$ over $Q=2$.

Similarly the element $\sqrt{3}\in R$ satisfies the polynomial x^2-3 over Q. Also x^2-3 is an irreducible polynomial over Q. Therefore degree of $\sqrt{3}$ over $Q=$ the degree of $x^2-3=2$. Also $[Q(\sqrt{3}):Q]=2$.

Let $\theta=\sqrt{2}+\sqrt{3} \Rightarrow \theta^2=5+2\sqrt{6}$

$\Rightarrow \theta^4=49+20\sqrt{6} \Rightarrow \theta^4=10\,(5+2\sqrt{6})-1$

$\Rightarrow \theta^4=10\theta^2-1 \Rightarrow \theta^4-10\theta^2+1=0$

$\Rightarrow \theta$ satisfies the polynomial x^4-10x^2+1 over Q.

(b) Since $\sqrt{2}+\sqrt{3} \in Q(\sqrt{2}, \sqrt{3})$, therefore

$$Q(\sqrt{2}+\sqrt{3}) \subseteq Q(\sqrt{2}, \sqrt{3}).$$

We shall now prove the converse. Since $Q(\sqrt{2}+\sqrt{3})$ is a field, therefore $(\sqrt{2}+\sqrt{3})^3=11\sqrt{2}+9\sqrt{3} \in Q(\sqrt{2}+\sqrt{3})$.

Also $-9\,(\sqrt{2}+\sqrt{3}) \in Q(\sqrt{2}+\sqrt{3})$.

$\therefore\quad \frac{1}{2}\,[(11\sqrt{2}+9\sqrt{3})+(-9)\,(\sqrt{2}+\sqrt{3})]=\sqrt{2} \in Q(\sqrt{2}+\sqrt{3})$.

$\therefore\quad \sqrt{2}+\sqrt{3}-\sqrt{2}=\sqrt{3} \in Q(\sqrt{2}+\sqrt{3})$.

Thus both $\sqrt{2}, \sqrt{3} \in Q(\sqrt{2}+\sqrt{3})$.

$\therefore\quad Q(\sqrt{2}, \sqrt{3}) \subseteq Q(\sqrt{2}+\sqrt{3})$.

Hence $Q(\sqrt{2}, \sqrt{3})=Q(\sqrt{2}+\sqrt{3})$.

(b) Let $L=Q(\sqrt{2})$. Then $[L : Q]=2$.

Also x^2-3 is an irreducible polynomial over L satisfied by $\sqrt{3}$. Therefore $[L(\sqrt{3}) : L]=2$.

Now $[L\,(\sqrt{3}) : Q]=[L\,(\sqrt{3}) : L]\,[L : Q]=2 \times 2=4$.

But $L(\sqrt{3})=(Q(\sqrt{2}))\,(\sqrt{3})=Q(\sqrt{2}, \sqrt{3})=Q(\sqrt{2}+\sqrt{3})$.

$\therefore\quad [Q(\sqrt{2}+\sqrt{3}) : Q]=4$.

This implies that $\sqrt{2}+\sqrt{3}$ is of degree 4 over Q.

Ex. 4. *Let a field L be a finite extension of a field K. Define the degree $[L : K]$ of L over K. Let Q denote the field of rational numbers, $K=Q(\sqrt{2})$, $L=K(\sqrt{3})$. Prove that $[L : K]=2$ and $[K : Q]=2$. What do you conclude about $[L : Q]$? Prove the theorem which you use for drawing your conclusion.* **(Gujrat 1970)**

Solution. Proceed as in Ex. 3.

Ex. 5. *If $a, b \in K$ are algebraic over F of degrees m and n, respectively, and if m and n are relatively prime, prove that $F(a, b)$ is of degree mn over F.*

Solution. Let $[F(a, b) : F]=k$.

We have $k=[F(a, b) : F]=[F(a, b) : F(a\,]\,[F(a) : F]$
$=[F\,(a, b) : F\,(a)]\,m$, since a is algebraic of degree m over
$F \Rightarrow [F(a) : F]=m$.

$\therefore\quad m$ is a divisor of k.

Similarly n is a divisor of k.

Now $[F\,(a, b) : F] \leqslant mn$.

[See the proof of theorem 12, page 492]

$\therefore\quad k \leqslant mn$.

Since $m \mid k$ and $n \mid k$ and m and n are relatively prime, therefore $mn \mid k$. So we cannot have $k<mn$. Hence $k=mn$.

Ex. 6. *Every finite extension K of a field F is algebraic and may be obtained from F by the adjunction of finitely many algebraic elements.*

Solution. If $[K:F]=1$, then $K=F$ and the result is trivial. So let $[K:F]=n>1$.

Let a be an arbitrary element in K but not in F. Since K is a field and $a \in K$, therefore the $n+1$ elements $1, a, a^2, \ldots, a^{n-1}, a^n$ are all in K. Since the dimension of the vector space $K(F)$ is n, therefore these $n+1$ elements of K are linearly dependent over F. So there exist elements of $\alpha_0, \alpha_1, \ldots, \alpha_n \in F$, not all 0, such that

$$\alpha_0 1+\alpha_1 a+\alpha_2 a^2+\ldots+\alpha_n a^n=0$$

$$\Rightarrow a \text{ is algebraic over } F.$$

Therefore K is an algebraic extension of F.

Now $F \subset F(a) \subseteq K$.

Also $n=[K:F]=[K:F(a)]\,[F(a):F]$. ...(1)

If $[F(a):F]=n$, then $[K:F(a)]=1$. This implies that $K=F(a)$ and our result is proved.

If $[F(a):F]\neq n$, then suppose that $[F(a):F]=m<n$.

Since $a \notin F$, therefore $m>1$. Also from (1), $[K:F(a)]>1$. Take an element b in K but not in $F(a)$. Then we can show as above that b is algebraic over F. Since a and b are algebraic over F, therefore $F(a, b)$ is a finite extension of F.

Let $[F(a, b):F]=p$ where $p>m$ because $F(a)$ is a proper subset of $F(a, b)$.

Now $n=[K:F]=[K:F(a, b)]\,[F(a, b):F]=[K:F(a, b)]\,p$.

If $p=n$, then $F(a, b)=K$ and our result is proved. If $p\neq n$, then we may continue the above process a finite number of times till we get $[F(a, b, \ldots, k):F]=n$. Then

$$n=[K:F]=[K:F(a, b, \ldots, k)]\,[F(a, b, \ldots, k):F]$$
$$=[K:F(a, b, \ldots, k)]\,n.$$

$\therefore\ [K:F(a, b, \ldots, k)]=1$

$\Rightarrow K=F(a, b, \ldots, k)$ where $a, b, \ldots, k$ are algebraic over F. This proves the result.

Ex. 7. *Let K be an extension of a field F and let $a \in K$ be algebraic over F. Then $F(a)$ is isomorphic to $F[x]/V$ where V is the ideal of $F[x]$ generated by the minimal polynomial for a over F.*

(Banaras 1971, 72)

Solution. Let $p(x)$ be the minimal polynomial for a over F. Then $p(x)$ is irreducible over F *i.e.*, $p(x)$ is a prime element of $F[x]$. Let $V=\{h(x) \in F[x] : h(a)=0\}$.

Then V is an ideal of $F[x]$. Now $F[x]$ is a principal ideal ring and $p(x)$ is an element of lowest degree in the ideal V of $F[x]$. Therefore V is the ideal of $F[x]$ generated by $p(x)$. Since $p(x)$ is irreducible, therefore V is a maximal ideal of $F[x]$. Consequently $F[x]/V$ is a field.

Let ψ be a mapping from $F[x]$ into $F(a)$ defined as follows :

$$f(x)\psi=f(a) \text{ for any } f(x)\in F[x].$$

Obviously ψ is a homomorphic mapping from the ring $F[x]$ into the field $F(a)$. The kernel of ψ is nothing but the ideal V of $F[x]$. Therefore by the fundamental theorem on ring homomorphism, $F[x]/V$ is isomorphic to $F[x]\psi$ which is the image of $F[x]$ under ψ. Now $F[x]\psi$ is a subset of $F(a)$. Since $F[x]/V$ is a field, therefore $F[x]\psi$ is a subfiled of $F(a)$. Now $x\in F[x]$ and by the definition of ψ, we have $x\psi=a$. Therefore $a\in F[x]\psi$. Also if $\alpha\in F$, then $\alpha\in F[x]$. By the definition of ψ, we have $\alpha\psi=\alpha$. Therefore $\alpha\in F[x]\psi$. Thus $F[x]\psi$ is a subfield of $F(a)$ and it contains both F and a. But $F(a)$ is the smallest subfield of K containing both F and a. Therefore we must have $F[x]\psi=F(a)$. Hence $F[x]/V$ is isomorphic to $F(a)$.

Note. If a is algebraic of degree n over F, then $p(x)$ is of degree n. Therefore the dimension of $F[x]/V$, as a vector space over F is n. By virtue of isomorphism between $F[x]/V$ and $F(a)$ we conclude that the dimension of $F(a)$, as a vector space over F is also n. Thus $[F(a):F]=n$. In this way we get an alternative proof for theorem 8.

§ 5. **Roots of Polynomials.** Let F be any field and let $p(x)$ be any polynomial in $F[x]$. Our aim is now to find a field K which is an extension of F and in which $p(x)$ has a root.

Roots of a polynomial. Definition. *Let F be any field and let $p(x)\in F[x]$. Then an element a lying in some extension field of F is called a root of $p(x)$ if $p(a)=0$.*

Theorem 14. Remainder Theorem.

If $p(x)\in F[x]$ and if K is an extension of F, then for any element $c\in K$, $p(x)=(x-c)\,q(x)+p(c)$ where $q(x)\in K[x]$ and where deg $q(x)=$ deg $p(x)-1$. **(Kanpur 1980; Meerut 73; Vikram 76)**

Proof. We have $F\subseteq K$

$$\Rightarrow F[x]\subseteq K[x]$$

$$\Rightarrow p(x)\in K[x] \qquad [\because p(x)\in F[x]]$$

Now the polynomials $p(x)$ and $x-c$ are both in $K[x]$. Therefore by division algorithm there exist polynomials $q(x)$ and $r(x)$ in $K[x]$ such that

$$p(x)=(x-c)\ q(x)+r(x),$$

where either $r(x)=0$ or deg $r(x)$ is less than the degree of $x-c$. But the degree of $x-c$ is 1. Therefore either $r(x)=0$ or deg $r(x)=0$. Hence $r(x)$ is a constant polynomial in $K[x]$ *i.e.*, $r(x)$ is simply an element, say r, in K. Thus

$$p(x)=(x-c)q(x)+r$$
$$\Rightarrow p(c)=(c-c)\ q(c)+r \quad \text{[Putting } x=c \text{ on both sides]}$$
$$\Rightarrow p(c)=0\ q(c)+r \Rightarrow p(c)=r.$$

Therefore $p(x)=(x-c)\ q(x)+p(c)$ ...(1)

Now suppose deg $p(x)=n$ and deg $q(x)=m$. The degree of the polynomial on the right hand side of (1) is then $m+1$. By the definition of equality of two polynomials we must have

$$n=m+1 \Rightarrow m=n-1 \Rightarrow \deg q(x)=\deg p(x)-1.$$

Corollary. Factor Theorem. *If* $a \in K$ *is a root of* $p(x) \in F[x]$, *where* $F \subseteq K$, *then in* $K[x]$, $(x-a) \mid p(x)$. **(Kanpur 1980)**

Proof. Let $p(x) \in F[x]$ and let $a \in K$ where K is an extension field of F. Then by remainder theorem in $K[x]$, we have

$$p(x)=(x-a)\ q(x)+p(a)$$
$$=(x-a)\ q(x)+0 \quad [\because\ a \text{ is a root of } p(x) \Rightarrow p(a)=0]$$
$$=(x-a)\ q(x).$$

Therefore in $K[x]$ we have $x-a$ is a divisor of $p(x)$. Thus

$$(x-a) \mid p(x) \text{ in } K[x].$$

Multiple root. Definition. *Let F be any field and let* $p(x) \in F[x]$. *If K is any extension of F, then* $a \in K$ *is said to be a root of* $p(x)$ *of multiplicity m if in* $K[x]$, *we have*

$(x-a)^m$ *is a divisor of* $p(x)$ *whereas* $(x-a)^{m+1}$ *is not a divisor of* $p(x)$.

A root of multiplicity 1 is called a **simple root** and a root of multiplicity > 1 is called a **multiple root**.

Now we are going to face an important problem. Suppose $p(x) \in F[x]$ and K is any extension field of F. The problem is that how many roots can $p(x)$ have in K. If a is a root of $p(x)$ in K of multiplicity m, then for this counting purpose we shall count it as m roots and not as one root.

Theorem 15. *A polynomial of degree n over a field can have at most n roots in any extension field.*

(Kanpur 1969; Banaras 72; Meerut 74)

Proof. We shall prove the theorem by induction on n, the degree of the polynomial $p(x)$

To start the induction let $p(x)$ be a polynomial of degree one over any field F. Let $p(x)=a_0x+a_1$ where $a_0, a_1 \in F$ and $a_0 \neq 0$. Let a be a root of $p(x)$ in some extension field of F. Then $p(a)=a_0a+a_1=0$. This gives $a=-\dfrac{a_1}{a_0}$ which is a unique element of F. Thus in this case $p(x)$ has the unique root $-\dfrac{a_1}{a_0}$ *i.e.*, $p(x)$ has one and exactly one root in any extension field of F. In this way the theorem is true when $p(x)$ is of degree 1.

Now assume as our induction hypothesis that the theorem is true in any field for all polynomials of degree less than n. Let $p(x)$ be a polynomial of degree n over a field F. Let K be any extension field of F. If $p(x)$ has no roots in K, then the theorem is obviously true because then the number of roots of $p(x)$ in K is zero which is definitely at most n. So let us suppose that $p(x)$ has at least one root, say, $a \in K$. Let a be a root of multiplicity m. Then in $K[x]$, we have

$$(x-a)^m \text{ is a divisor of } p(x)$$

$$\Rightarrow \deg (x-a)^m \leqslant \deg p(x) \Rightarrow m \leqslant n.$$

Since in $K[x]$ we have $(x-a)^m$ is a divisor of $p(x)$, therefore let $p(x)=(x-a)^m q(x)$ where $q(x) \in K[x]$.

We have $\deg q(x)=\deg p(x)-\deg (x-a)^m=n-m$ which is definitely less than n because $1 \leqslant m \leqslant n$.

Now a is a root of $p(x)$ of multiplicity m. Therefore $(x-a)^{m+1}$ is not a divisor of $p(x)=(x-a)^m q(x)$. This implies that $(x-a)$ is not a divisor of $q(x)$. Therefore a is not a root of $q(x)$. [See corollary to theorem 14]. Now let $b \neq a$ be a root of $p(x)$ in K. Then on putting $x=b$ in $p(x)=(x-a)^m q(x)$, we get

$$0=p(b)=(b-a)^m q(b).$$

Since K is a field and $0 \neq (b-a)^m \in K$ and $q(b) \in K$, therefore we must have

$$q(b)=0 \Rightarrow b \text{ is a root of } q(x) \text{ in } K.$$

Thus any root of $p(x)$, in K, other than a must also be a root of $q(x)$ in K.

Now $q(x)$ is of degree $n-m$ which is less than n. Therefore by our induction hypothesis $q(x)$ has at most $n-m$ roots in K and none of these roots is equal to a because a is not a root of $q(x)$. Thus $q(x)$ has at most $n-m$ roots other than a in K

$\Rightarrow p(x)$ has at most $n-m$ roots other than a in K

$\Rightarrow p(x)$ has at most $(n-m)+m=n$ roots in K, the root a of $p(x)$ of multiplicity m being counted m times.

The proof is now complete by induction.

Theorem 16. *If $p(x)$ is a polynomial in $F[x]$ of degree $n \geqslant 1$ and is irreducible over F, then there is an extension E of F, such that $[E:F]=n$, in which $p(x)$ has a root.*

(I. A. S. 1972; Banaras 70; Meerut 79, 83 P)

Proof Let $F[x]$ be the ring of polynomials over F and let $V=(p(x))$ be the ideal of $F[x]$ generated by $p(x)$. Since $p(x)$ is irreducible over F, therefore V is a maximal ideal of $F[x]$. Consequently $E=F[x]/V$ is a field. We shall show that the field E satisfies the conclusions of the theorem.

First we shall show that E can be regarded as an extension of F even though E does not contain the elements of F in their original form. For this we shall show that the field F can be imbedded in the field E.

Let ψ be the mapping from F into E defined by

$$\psi(\alpha)=V+\alpha \ \forall \alpha \in F.$$

ψ **is one-one.** Let $\alpha, \beta \in F$. We have

$\psi(\alpha)=\psi(\beta) \Rightarrow V+\alpha=V+\beta \Rightarrow \alpha-\beta \in V$

$\Rightarrow \alpha-\beta=f(x)\,p(x)$ for some $f(x) \in F[x]$

[Note that V is the ideal generated by $p(x)$]

$\Rightarrow f(x)=0$ because if $f(x) \neq 0$, then the polynomial $f(x)\,p(x)$ is of positive degree and so it cannot be equal to the constant polynomial $\alpha-\beta$

$\Rightarrow \alpha-\beta=0 \Rightarrow \alpha=\beta$.

$\therefore$ ψ is one-one.

Also $\psi(\alpha+\beta)=V+(\alpha+\beta)=(V+\alpha)+(V+\beta)=\psi(\alpha)+\psi(\beta)$, and $\psi(\alpha\beta)=V+\alpha\beta=(V+\alpha)(V+\beta)=\psi(\alpha)\psi(\beta)$.

Thus ψ is an isomorphism from F into E. Let F^* be the image of F into E under the mapping ψ *i.e.*, let $F^*=\{V+\alpha : \alpha \in F\}$. Then ψ is an isomorphism of F onto F^* and F^* is a subfield of E isomorphic to F. If we identify F and F^* *i.e.*, if we identify $\alpha \in F$ with $V+\alpha \in F^*$, then E can be regarded as an extension of F.

Now we claim that E is a finite extension of F. For this we shall prove that the n elements $V+1, V+x, V+x^2, \ldots, V+x^{n-1}$ form a basis of E over F. [For proof see Ex. 2 page 494].

$\therefore$ $[E:F]=n$.

Finally we shall show that $p(x)$ has a root in E.

Let $p(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n$ where $a_0, a_1, \ldots, a_n \in F$.

First let us make $p(x)$ as a polynomial over E with the help of the identification we have made between F and F^*. So let us replace a_0 by $V+a_0$, a_1 by $V+a_1, \ldots, a_n$ by $V+a_n$. Then

$$p(x)=(V+a_0)+(V+a_1)\, x+\cdots+(V+a_n)\, x^n.$$

We shall show that $V+x \in E$ satisfies $p(x)$. We have

$$p(V+x)=(V+a_0)+(V+a_1)\,(V+x)+(V+a_2)\,(V+x)^2 +\cdots+(V+a_n)\,(V+x)^n$$

$$=(V+a_0)+(V+a_1)\,(V+x)+(V+a_2)\,(V+x^2)+\ldots+(V+a_n)\,(V+x^n)$$

[Note that $(V+x)^2=(V+x)\,(V+x)=V+x^2$, and so on]

$$=(V+a_0)+(V+a_1x)+(V+a_2x^2)+\ldots+(V+a_nx^n)$$

[by def. of multiplication of cosets we have $[V+f(x)][V+g(x)]=V+f(x)\, g(x)]$

$$=V+a_0+a_1x+a_2x^2+\ldots+a_nx^n$$

[by def. of addition of cosets we have $[V+f(x)]+[V+g(x)]=V+f(x)+g(x)]$

$=V+p(x)$

$=V$, since $p(x) \in V$

$=$ the zero element of the field E.

[Note that the zero element of the field $F[x]/V$ is nothing but the coset V itself].

Thus $V+x$ satisfies $p(x)$. Therefore $V+x$ is a root of $p(x)$.

The proof of the theorem is now complete.

Corollary. *If $f(x) \in F[x]$, then there is a finite extension E of F in which $f(x)$ has a root. Moreover, $[E : F] \leqslant \deg f(x)$.*

(Meerut 1973; Banaras 70)

Proof. Let $p(x)$ be an irreducible factor of $f(x)$. Let

$$f(x)=p(x)\; q(x).$$

We have $\deg p(x) \leqslant \deg f(x)$.

Let a be the root of $p(x)$ in some extension field K of F. Then $p(a)=0$. We have

$$f(a)=p(a)\; q(a)=0\; q(a)=0.$$

Thus any root of $p(x)$ in some extension field of F is also a root of $f(x)$ in that extension field.

Since $p(x)$ is irreducible over F, therefore by the above theorem there is an extension E of F with

$$[E : F]=\deg p(x) \leqslant \deg f(x)$$

in which $p(x)$ has a root.

Hence there is a finite extension E of F with $[E:F] \leqslant \deg f(x)$ in which $f(x)$ has a root.

Theorem 17. *Let $f(x) \in F[x]$ be of degree $n \geqslant 1$. Then there is a finite extension E of F of degree at most $n\,!$ in which $f(x)$ has n roots (and so, a full complement of roots).*

(I.A.S. 1973; Meerut 74)

Proof. We shall prove the theorem by induction on n, the degree of $f(x)$.

To start the induction let $f(x) \in F[x]$ be of degree 1. Let $f(x) = a_0x + a_1$ where $a_0, a_1 \in F$ and $a_0 \neq 0$. Now F itself is an extension of F and $[F:F]=1$. Also $-\dfrac{a_1}{a_0} \in F$ is a root of a_0x+a_1. Thus if $\deg f(x)=1$, then there is a finite extension F of F of degree at most $1\,!=1$ in which $f(x)$ has one root.

Now assume as our induction hypothesis that the theorem is true in any field for all polynomials of degree less than n. Let $f(x)$ be a polynomial of degree n over a field F. By corollary to theorem 16, there is an extension E_0 of F with $[E_0 : F] \leqslant n$ in which $f(x)$ has a root, say, α. By factor theorem in $E_0[x]$, $f(x)$ factors as

$$f(x)=(x-\alpha)\, q(x)$$

where $\qquad \deg q(x) = \deg f(x) - 1 = n-1.$

Now $q(x)$ is a polynomial over E_0 of degree $n-1$. Since $\deg q(x)$ is less than n, therefore by our induction hypothesis there is an extension E of E_0 degree at most $(n-1)\,!$ in which $q(x)$ has $n-1$ roots Since any root of $f(x)$ is either α or a root of $q(x)$, therefore we obtain in E all n roots of $f(x)$.

Now E is an extension of E_0 and E_0 is an extension of F implies that E is an extension of F. We have

$$[E:F]=[E:E_0]\,[E_0:F]$$
$$\leqslant (n-1)\,!\,n = n\,!.$$

Thus E is a finite extension of F of degree at most $n\,!$ in which $f(x)$ has n roots.

The proof of the theorem is now complete by induction.

Splitting field or Decomposition field. Definition.

If $f(x) \in F[x]$, a finite extension E of F is said to be a splitting field over F for $f(x)$ if over E (that is, in $E[x]$), but not over any proper subfield of E, $f(x)$ can be factored as a product of linear (first degree) factors. **(Banaras 1971; Meerut 78, 79, 81, 84 P)**

The field F is called the **base field** or the **initial field.**

Theorem 18. *There exists a splitting field for every $f(x) \in F[x]$.*
(Delhi 1970; Aligarh 66, Banaras 71, Meerut 74)

Proof. Let $f(x) \in F[x]$ be of degree n. First we shall prove that there exists a finite extension E of F of degree at most n ! in which $f(x)$ has n roots. [For proof see theorem 17].

Let $f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n$, $a_0 \neq 0$.

Let $\alpha_1, \ldots, \alpha_n$ be the n roots in E of $f(x)$. Then by factor theorem $f(x)$ can be factored over E as

$$f(x)=a_0(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n).$$

In this way $f(x)$ splits up completely over E as a product of first degree factors. Thus we see that there exists a finite extension E of F which decomposes $f(x)$ as a product of linear factors. Consequently a finite extension of F of minimal degree exists which also possesses this property. This minimal extension will be a splitting field for $f(x)$ because no proper subfield of this minimal extension can split $f(x)$ as a product of linear factors.

Another way of defining the splitting field.

An extension E of a field F is said to be a splitting field of

$$f(x) \in F[x], \text{ if } f(x) \in E[x]$$

is expressible as

$$f(x)=a_0(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n),$$

where $\quad a_0 \in F, \alpha_1, \ldots, \alpha_n \in E,$

and $\quad E=F(\alpha_1, \alpha_2, \ldots, \alpha_n).$

Uniqueness of the splitting field

Now we shall show that the splitting field of a polynomial is unique apart from isomorphism. Let E_1 and E_2 be two splitting fields of $f(x) \in F[x]$ Let

$$f(x)=a(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n) \text{ over } E_1$$

and $\quad f(x)=a(x-\beta_1)(x-\beta_2)\ldots(x-\beta_n)$ over E_2.

We shall show that the fields

$$F(\alpha_1, \ldots, \alpha_n) \text{ and } F(\beta_1, \ldots, \beta_n)$$

are isomorphic by an isomorphism leaving every element of F fixed. Before proving this main theorem we shall first prove some pre-requisite results.

Continuation of an isomorphic mapping. Definition.

Let F and F′ be two isomorphic fields and let E and E′ be extension fields of F and F′ respectively. An isomorphism $\sigma : E \leftrightarrow E'$ *is called a continuation of the isomorphism* $\psi : F \leftrightarrow F'$ *if* $\sigma(\alpha)=\psi(\alpha)$ *for all* α *in F.*

Let ψ be an isomorphism of a field F onto a field F'. For the sake of convenience we shall denote the image of any $\alpha \in F$ under ψ by α'. Thus $\alpha\psi = \alpha'$. Here by $\alpha\psi$ we mean $\psi(\alpha)$. Note that $\alpha\psi$ is another way of writing the image of α under ψ.

Theorem 19. *Let ψ be an isomorphism of a field F onto a field F' such that $\alpha\psi = \alpha'$ for every $\alpha \in F$. Show that there is an isomorphism ψ^* of $F[x]$ onto $F'[t]$ with the property that*

$$\alpha\psi^* = \alpha\psi = \alpha' \text{ for every } \alpha \in F.$$

Proof. ψ is an isomorphism of a field F onto a field F'.

For any $\alpha \in F$ we write $\alpha\psi = \alpha'$. Let us define a mapping ψ^* from $F[x]$ into $F'[t]$ as follows:

For an arbitrary polynomial $f(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n \in F[x]$ we define ψ^* by

$$\begin{aligned} f(x)\,\psi^* &= (\alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n)\,\psi^* \\ &= \alpha_0\psi + \alpha_1\psi t + \ldots + \alpha_n\psi t^n \\ &= \alpha_0' + \alpha_1' t + \ldots + \alpha_n' t^n = f'(t), \text{ say.} \end{aligned}$$

We shall show that the mapping ψ^* satisfies the conclusions of the theorem.

ψ^* is one-one.

Let $f(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n$, $g(x) = \beta_0 + \beta_1 x + \ldots + \beta_m x^m$ be any two elements of $F[x]$. We have

$$f(x)\,\psi^* = g(x)\,\psi^*$$

$\Rightarrow (\alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n)\,\psi^* = (\beta_0 + \beta_1 x + \ldots + \beta_m x^m)\,\psi^*$

$\Rightarrow \alpha_0' + \alpha_1' t + \ldots + \alpha_n' t^n = \beta_0' + \beta_1' t + \ldots + \beta_m' t^m$

$\Rightarrow n = m$ and $\alpha_i' = \beta_i'$ for each $i = 0, 1, \ldots, n$

$\Rightarrow n = m$ and $\alpha_i\psi = \beta_i\psi$ for each i

$\Rightarrow n = m$ and $\alpha_i = \beta_i$ for each i $\qquad [\because \psi$ is one-one$]$

$\Rightarrow f(x) = g(x) \Rightarrow \psi^*$ is one-one.

ψ^* is onto.

Let $\delta_0' + \delta_1' t + \ldots + \delta_n' t^n$ be any element of $F'[t]$ where $\delta_0', \delta_1', \ldots, \delta_n' \in F'$. Since ψ is onto F', therefore $\exists\ \delta_0, \delta_1, \ldots, \delta_n \in F$ such that $\delta_0\psi = \delta_0'$, $\delta_1\psi = \delta_1', \ldots, \delta_n\psi = \delta_n'$. Now

$$\delta_0 + \delta_1 x + \ldots + \delta_n x^n \in F[x]$$

and we have $(\delta_0 + \delta_1 x + \ldots + \delta_n x^n)\,\psi^* = \delta_0' + \delta_1' t + \ldots + \delta_n' t^n$.

$\therefore$ the mapping ψ^* is onto.

ψ^* preserves additions.

Let $\qquad f(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n$, $g(x) = \beta_0 + \beta_1 x + \ldots + \beta_m x^m$ be any two elements of $F[x]$. Without loss of generality let $n \geqslant m$. First let us take the case when $n > m$.

We have

$$\begin{aligned}[f(x)+g(x)]\,\psi^* &= [(\alpha_0+\beta_0)+(\alpha_1+\beta_1)\,x+\dots+(\alpha_m+\beta_m)\,x^m \\ &\qquad\qquad +\alpha_{m+1}x^{m+1}+\dots+\alpha_n x^n]\,\psi^* \\ &=(\alpha_0+\beta_0)\,\psi+(\alpha_1+\beta_1)\,\psi t+\dots+(\alpha_m+\beta_m)\,\psi t^m \\ &\qquad\qquad +\alpha_{m+1}\psi t^{m+1}+\dots+\alpha_n\psi t^n \\ &=(\alpha_0\psi+\beta_0\psi)+(\alpha_1\psi+\beta_1\psi)\,t+\dots+(\alpha_m\psi+\beta_m\psi)\,t^m \\ &\qquad\qquad +\alpha_{m+1}\psi t^{m+1}+\dots+\alpha_n\psi t^n \\ &\qquad\qquad [\because\ \psi \text{ is an isomorphism}] \\ &=(\alpha_0'+\beta_0')+(\alpha_1'+\beta_1')\,t+\dots+(\alpha_m'+\beta_m')\,t^m \\ &\qquad\qquad +\alpha'_{m+1}t^{m+1}+\dots+\alpha_n't^n \\ &=(\alpha_0'+\alpha_1't+\dots+\alpha_m't^m+\alpha'_{m+1}t^{m+1}+\dots+\alpha_n't^n) \\ &\qquad\qquad +(\beta_0'+\beta_1't+\dots+\beta_m't^m) \\ &=f(x)\,\psi^*+g(x)\,\psi^*.\end{aligned}$$

Similarly when $n=m$, we can show that

$$[f(x)+g(x)]\,\psi^*=f(x)\,\psi^*+g(x)\,\psi^*.$$

ψ^* preserves multiplications. We have

$$\begin{aligned}[f(x)\,g(x)]\,\psi^* &= [(\alpha_0+\alpha_1x+\dots+\alpha_nx^n)\,(\beta_0+\beta_1x+\dots+\beta_mx^m)]\,\psi^* \\ &=[\alpha_0\beta_0+(\alpha_0\beta_1+\alpha_1\beta_0)\,x+\dots+\alpha_n\beta_mx^{n+m}]\,\psi^* \\ &=(\alpha_0\beta_0)\,\psi+(\alpha_0\beta_1+\alpha_1\beta_0)\,\psi t+\dots+(\alpha_n\beta_m)\,\psi t^{n+m} \\ &=\alpha_0'\beta_0'+(\alpha_0'\beta_1'+\alpha_1'\beta_0')\,t+\dots+(\alpha_n'\beta_m')\,t^{n+m} \\ &\qquad [\because\ \psi \text{ preserves additions and multiplications} \\ &\qquad\qquad i.e.,\ (\alpha+\beta)\,\psi=\alpha\psi+\beta\psi=\alpha'+\beta' \text{ and} \\ &\qquad\qquad (\alpha\beta)\,\psi=(\alpha\psi)\,(\beta\psi)=\alpha'\beta'] \\ &=(\alpha_0'+\alpha_1'\,t+\dots+\alpha_n't^n)\,(\beta_0'+\beta_1't+\dots+\beta_m't^m) \\ &=[f(x)\,\psi^*]\,[g(x)\,\psi^*].\end{aligned}$$

Hence ψ^* is an isomorphism of $F[x]$ onto $F'[t]$.

Further if $f(x)\in F[x]$ be simply taken as α where $\alpha\in F$, then by definition of ψ^*, we have

$$\alpha\psi^*=\alpha\psi=\alpha'.$$

Note. Form the above theorem we conclude that factorizations of $f(x)$ in $F[x]$ result in like factorizations of $f(x)\,\psi^*=f'(t)$ in $F'[t]$ and vice versa. In particular, $f(x)$ is irreducible in $F[x]$ if and only if $f'(t)$ is irreducible in $F'[t]$.

Theorem 20. *Let ψ be an isomorphism of a field F onto a field F' defined as $\alpha\psi=\alpha'$ for every $\alpha\in F$. For an arbitrary polynomial $f(x)=\alpha_0+\alpha_1x+\dots+\alpha_nx^n\in F[x]$ let us define $f'(t)=\alpha_0'+\alpha_1't+\dots+\alpha_n't^n\in F'[t]$. If $f(x)$ is irreducible in $F[x]$ show that there is an isomorphism θ of $F[x]/(f(x))$ onto $F'[t]/(f'(t))$ with the property that for every $\alpha\in F$, $\alpha\theta=\alpha\psi=\alpha'$.*

Proof. Let ψ^* be a function from $F[x]$ into $F'[t]$ defined by the formula

$$f(x)\,\psi^* = f'(t) \text{ for every } f(x) \in F[x].$$

Then by theorem 19 the mapping ψ^* is an isomorphism of $F[x]$ onto $F'[t]$.

Let $f(x)$ be irreducible in $F[x]$. Then $f'(t)$ will be irreducible in $F'[t]$. Let $V = (f(x))$ be the ideal of $F[x]$ generated by $f(x)$ and $V' = (f'(t))$ be the ideal of $F'[t]$ generated by $f'(t)$. Both V and V' are maximal ideals because $f(x)$ and $f'(t)$ are irreducible. Therefore $F[x]/V$ and $F'[t]/V'$ are both fields.

Let us define a mapping θ from $F[x]/V$ into $F'[t]/V'$ by the formula

$[V+g(x)]\,\theta = V'+g(x)\,\psi^* = V'+g'(t)$ for every $g(x) \in F[x]$.

The mapping θ is **well-defined.**

For this we are to show that if $V+g(x)=V+h(x)$, then

$$[V+g(x)]\,\theta = [V+h(x)]\,\theta \text{ where } g(x), h(x) \in F[x].$$

We have $V+g(x)=V+h(x)$

$\Rightarrow g(x)-h(x) \in V$

$\Rightarrow g(x)-h(x)=k(x)\,f(x)$ for some $k(x)\in F[x]$

$\Rightarrow [g(x)-h(x)]\,\psi^* = [k(x)\,f(x)]\,\psi^*$

$\Rightarrow g(x)\,\psi^* - h(x)\,\psi^* = [k(x)\,\psi^*]\,[f(x)\,\psi^*]$

[∵ ψ^* is an isomorphism]

$\Rightarrow g'(t)-h'(t)=k'(t)\,f'(t)$

$\Rightarrow g'(t)-h'(t)\in V' \Rightarrow V'+g'(t)=V'+h'(t)$

$\Rightarrow [V+g(x)]\,\theta=[V+h(x)]\,\theta.$

Therefore the mapping θ is well defined.

The mapping θ is **one-one.**

Let $g(x), h(x) \in F[x]$. Then

$[V+g(x)]\,\theta=[V+h(x)]\,\theta$

$\Rightarrow V'+g'(t)=V'+h'(t) \Rightarrow g'(t)-h'(t)\in V'$

$\Rightarrow g'(t)-h'(t)=k'(t)\,f'(t)$ for some $k'(t)\in F'[t]$

$\Rightarrow g(x)\,\psi^*-h(x)\,\psi^*=[k(x)\,\psi^*]\,[f(x)\,\psi^*]$

$\Rightarrow [g(x)-h(x)]\,\psi^*=[k(x)\,f(x)]\,\psi^*$

$\Rightarrow g(x)-h(x)=k(x)\,f(x)$ [∵ ψ^* is one-one]

$\Rightarrow g(x)-h(x)\in V \Rightarrow V+g(x)=V+h(x) \Rightarrow \theta$ is one-one.

The mapping θ is **onto.**

Since the mapping ψ^* is onto, therefore corresponding to any polynomial $g'(t)$ in $F'[t]$, we have a polynomial $g(x)$ in $F[x]$. Therefore $V'+g'(t)\in F'[t]/V' \Rightarrow \exists\; V+g(x)\in F[x]/V$ such that

$$[V+g(x)]\,\theta = V'+g'(t).$$

θ preserves additions and multiplications.

Let $g(x)$, $h(x) \in F[x]$. We have

$[\{V+g(x)\}+\{V+h(x)\}]\,\theta=[V+g(x)+h(x)]\,\theta$

$=V'+[g(x)+h(x)]\,\psi^*=V'+g(x)\,\psi^*+h(x)\,\psi^*$

$=V'+g'(t)+h'(t)=[V'+g'(t)]+[V'+h'(t)]$

$=[V+g(x)]\,\theta+[V+h(x)]\,\theta.$

Also $[\{V+g(x)\}\,\{V+h(x)\}]\,\theta=[V+g(x)\;h(x)]\,\theta$

$=V'+[g(x)\,h(x)]\;\psi^*=V'+[g(x)\,\psi^*]\;[h(x)\,\psi^*]$

$=V'+g'(t)\;h'(t)=[V'+g'(t)]\;[V'+h'(t)]$

$=[\{V+g(x)\}\,\theta]\;[\{V+h(x)\}\theta].$

Thus θ is an isomorphism of $F[x]/V$ onto $F'[t]/V'$.

In theorem 16 we have shown that the field F can be imbedded in the field $F[x]/V$ by identifying the element $\alpha \in F$ with the residue class (coset) $V+\alpha$ in $F[x]/V$. Similarly we can consider F' to be contained in $F'[t]/V'$. With this identification, for any $\alpha \in F$, we have

$$\begin{aligned}\alpha\theta&=(V+\alpha)\,\theta && [\because\ \alpha \text{ has been identified with } V+\alpha]\\ &=V'+\alpha\psi^*=V'+\alpha'\\ &=\alpha'. && [\because\ \alpha' \text{ has been identified with } V'+\alpha']\end{aligned}$$

Theorem 21. *Let ψ be an isomorphism of a field F onto a field F' such that $\alpha\psi=\alpha'$ for every $\alpha \in F$. Let $f(x)=\alpha_0+\alpha_1x+\ldots+\alpha_nx^n$ be an irreducible polynomial in $F[x]$ which is mapped onto the (also irreducible) polynomial $f'(t)=\alpha_0'+\alpha_1't+\ldots+\alpha_n't^n$ in $F'[t]$. If v is a root of $f(x)$ in some extension field of F and w is a root of $f'(t)$ in some extension field of F', then the field $F(v)$ is isomorphic to the field $F'(w)$ by an isomorphism σ such that*

(1) $v\sigma=w$

(2) $\alpha\sigma=\sigma\psi=\alpha'$ *for every* $\alpha \in F$.

Note. The condition (2) may also be stated as that the isomorphism σ is a continuation of the isomorphism ψ.

Proof. ψ is an isomorphism of a field F onto a field F' such that $\alpha\psi=\alpha'$ for every $\alpha \in F$. The irreducible polynomial $f(x)=\alpha_0+\alpha_1x+\ldots+\alpha_nx^n$ in $F[x]$ is mapped onto the irreducible polynomial $f'(t)=\alpha_0'+\alpha_1't+\ldots+\alpha_n't^n$ in $F'[t]$. We have $\deg f(x)=n=\deg f'(t)$.

Let v be a root of $f(x)$ in some extension field of F. Then v is algebraic over F. Since $f(x)$ is irreducible in $F[x]$, therefore $f(x)$ is a minimal polynomial for v over F. Consequently by theorem 5, we have $F(v)=\{\beta_0+\beta_1v+\ldots+\beta_{n-1}v^{n-1} : \beta_0, \beta_1, \ldots, \beta_{n-1} \in F\}$.

Also the expression $\beta_0+\beta_1 v+\ldots+\beta_{n-1}v^{n-1}$ for an element of $F(v)$ is unique. Similarly

$$F'(w)=\{\delta_0'+\delta_1'w+\ldots+\delta'_{n-1}w^{n-1}:\delta_0',\delta_1',\ldots,\delta'_{n-1}\in F'\}.$$

Also the expression $\delta_0'+\delta_1'w+\ldots+\delta'_{n-1}w^{n-1}$ for an element of $F'(w)$ is unique. Let c be an arbitrary element of $F(v)$. Then c can be written as

$c=\beta_0+\beta_1 v+\ldots+\beta_{n-1}v^{n-1}$, where $\beta_0,\beta_1,\ldots,\beta_{n-1}$ are unique elements of F.

Let us define a mapping σ from $F(v)$ into $F'(w)$ by the formula

$$\begin{aligned}c\sigma&=(\beta_0+\beta_1 v+\ldots+\beta_{n-1}v^{n-1})\,\sigma\\&=\beta_0\psi+\beta_1\psi w+\ldots+\beta_{n-1}\psi w^{n-1}\\&=\beta_0'+\beta_1'w+\ldots+\beta'_{n-1}w^{n-1}\end{aligned}$$

Since the expression for c in the form $\beta_0+\beta_1 v+\ldots+\beta_{n-1}v^{n-1}$ is unique, therefore the mapping σ is well defined.

σ is one-one. Let $c=\beta_0+\beta_1 v+\ldots+\beta_{n-1}v^{n-1}$, $d=\gamma_0+\gamma_1 v+\ldots+\gamma_{n-1}v^{n-1}$ be any two elements of $F(v)$. Then $c\sigma=d\sigma$

$\Rightarrow (\beta_0+\beta_1 v+\ldots+\beta_{n-1}v^{n-1})\,\sigma=(\gamma_0+\gamma_1 v+\ldots+\gamma_{n-1}v^{n-1})\,\sigma$

$\Rightarrow \beta_0'+\beta_1'w+\ldots+\beta'_{n-1}w^{n-1}=\gamma_0'+\gamma_1'w+\ldots+\gamma'_{n-1}w^{n-1}$

$\Rightarrow \beta'_0=\gamma'_0,\ \beta_1'=\gamma_1',\ldots,\ \beta'_{n-1}=\gamma'_{n-1}$ [Since the expression for an element of $F'(w)$ in the form $\beta_0'+\beta_1'w+\ldots+\beta'_{n-1}w^{n-1}$ is unique]

$\Rightarrow \beta_0\psi=\gamma_0\psi,\ \beta_1\psi=\gamma_1\psi,\ldots,\ \beta_{n-1}\psi=\gamma_{n-1}\psi$

$\Rightarrow \beta_0=\gamma_0,\ \beta_1=\gamma_1,\ldots,\ \beta_{n-1}=\gamma_{n-1}$ [$\because$ ψ is one-one]

$\Rightarrow c=d\Rightarrow \sigma$ is one-one.

σ is onto. Let $\delta_0'+\delta_1'w+\ldots+\delta'_{n-1}w^{n-1}$ be any element of $F'(w)$ where $\delta_0',\delta_1',\ldots,\delta'_{n-1}\in F'$. Since ψ is onto F', therefore $\exists\ \delta_0,\delta_1,\ldots,\delta_{n-1}\in F$ such that $\delta_0\psi=\delta_0',\ \delta_1\psi=\delta_1',\ldots,\ \delta_{n-1}\psi=\delta'_{n-1}$. Now $\delta_0+\delta_1 v+\ldots+\delta_{n-1}v^{n-1}\in F(v)$ and we have

$$(\delta_0+\delta_1 v+\ldots+\delta_{n-1}v^{n-1})\,\sigma=\delta_0'+\delta_1'w+\ldots+\delta'_{n-1}w^{n-1}.$$

$\therefore$ σ is onto $F'(w)$.

σ preserves additions. We have

$$\begin{aligned}(c+d)\,\sigma&=[(\beta_0+\beta_1 v+\ldots+\beta_{n-1}v^{n-1})+(\gamma_0+\gamma_1 v+\ldots+\gamma_{n-1}v^{n-1})]\,\sigma\\&=[(\beta_0+\gamma_0)+(\beta_1+\gamma_1)\,v+\ldots+(\beta_{n-1}+\gamma_{n-1})\,v^{n-1}]\,\sigma\\&=(\beta_0+\gamma_0)\,\psi+(\beta_1+\gamma_1)\,\psi w+\ldots+(\beta_{n-1}+\gamma_{n-1})\,\psi w^{n-1}\\&=(\beta_0\psi+\gamma_0\psi)+\ldots+(\beta_{n-1}\psi+\gamma_{n-1}\psi)\,w^{n-1}\\&=(\beta_0'+\gamma_0')+\ldots+(\beta'_{n-1}+\gamma'_{n-1})\,w^{n-1}\\&=(\beta_0'+\beta_1'w+\ldots+\beta'_{n-1}w^{n-1})+(\gamma_0'+\gamma_1'w+\ldots+\gamma'_{n-1}w^{n-1})\\&=(\beta_0+\beta_1 v+\ldots+\beta_{n-1}v^{n-1})\,\sigma+(\gamma_0+\gamma_1 v+\ldots+\gamma_{n-1}v^{n-1})\,\sigma\\&=c\sigma+d\sigma.\end{aligned}$$

σ preserves multiplications.

We have

$$
\begin{aligned}
(cd)\,\sigma &= [(\beta_0+\beta_1 v+\dots+\beta_{n-1}v^{n-1})\,(\gamma_0+\gamma_1 v+\dots+\gamma_{n-1}v^{n-1})]\,\sigma \\
&= [\beta_0\gamma_0+(\beta_0\gamma_1+\beta_1\gamma_0)\,v+\dots+\beta_{n-1}\gamma_{n-1}\,v^{2n-2}]\,\sigma \\
&= (\beta_0\gamma_0)\,\psi+(\beta_0\gamma_1+\beta_1\gamma_0)\,\psi w+\dots+(\beta_{n-1}\gamma_{n-3})\,\psi w^{2n-2} \\
&= \beta_0'\gamma_0'+(\beta_0'\gamma'_1)+\beta_1'\gamma_0')\,w+\dots+(\beta'_{n-1}\gamma'_{n-1})\,w^{2n-2} \\
&\qquad [\because\ \psi \text{ preserves additions and multiplications}] \\
&= (\beta_0'+\beta_1'w+\dots+\beta'_{n-1}w^{n-1})\,(\gamma_0'+\gamma_1'w+\dots+\gamma'_{n-1}w^{n-1}) \\
&= [(\beta_0+\beta_1 v+\dots+\beta_{n-1}v^{n-1})\,\sigma]\,[(\gamma_0+\gamma_1 v+\dots+\gamma_{n-1}v^{n-1})\,\sigma], \\
&= (c\sigma)\,(d\sigma).
\end{aligned}
$$

Hence σ is an isomorphism of $F(v)$ onto $F'(w)$.

Now $v \in F(v)$ can be uniquely written as

$$v=0+1v+0v^2+\dots+0v^{n-1}.$$

$$\therefore \quad v\sigma=0+1w+0w^2+\dots+0w^{n-1}=w.$$

Also if $\alpha \in F$ then $\alpha \in F(v)$ can be uniquely written as

$$\alpha=\alpha+0v+0v^2+\dots+0v^{n-1}.$$

$$\therefore \quad \alpha\sigma=\alpha'+0w+0w^2+\dots+0w^{n-1}=\alpha'.$$

This completes the proof of the theorem.

Corollary. *If $f(x) \in F[x]$ is irreducible and if a, b are two roots of $f(x)$, then $F(a)$ is isomorphic to $F(b)$ by an isomorphism which takes a onto b and which leaves every element of F fixed.*

Proof. In the above theorem replace the field F' by F *i.e.*, take $F'=F$. Take the isomorphism ψ as the identity map of F *i.e.*, $\psi : F \to F$ such that $\alpha\psi=\alpha \ \forall \ \alpha \in F$. Then ψ leaves every element of F fixed and ψ is an isomorphism of F onto F.

Take $v=a$ and $w=b$.

Then $F(a)$ is isomorphic to $F(b)$ by an isomorphism σ such that

(1) $a\sigma=b$,

(2) $\alpha\sigma=\alpha\psi=\alpha$ for every $\alpha \in F$ *i.e.*, σ leaves every element of F fixed.

Now re-write the proof of the above theorem and thus complete the proof of this corollary because the corollary in itself is very important.

Theorem 22. *Let ψ be an isomorphism of a field F onto a field F' defined as $\alpha\psi=\alpha'$ for every $\alpha \in F$. Corresponding to a polynomial $f(x)=\alpha_0+\alpha_1 x+\dots+\alpha_n x^n$ in $F[x]$,*

let $f'(t)=\alpha_0'+\alpha_1't+\dots+\alpha_n't^n$

be a polynomial in $F'[t]$. Show that the splitting fields E and E' of $f(x) \in F[x]$ and $f'(t) \in F'[t]$, respectively, are isomorphic by an isomorphism ϕ with the property that $\alpha\phi=\alpha\psi=\alpha'$ for every $\alpha \in F$.

Proof. We shall prove the theorem by induction on the degree of some splitting field over the initial field.

To start the induction let $[E : F]=1$. Then $E=F$ and therefore $f(x)$ resolves into a product of linear factors over F itself. By theorem 19, $f'(t)$ must also resolve into a product of linear factors over F' itself. Therefore F' is a splitting field for $f'(t)$ *i.e.*, $F'=E'$. Hence $\phi=\psi$ provides us with an isomorphism of E onto E' coinciding with ψ on F.

Now assume as our induction hypothesis that the theorem is true for any field F_0 and any polynomial $g(x)\in F_0[x]$ provided the degree of some splitting field E_0 of $g(x)$ over the initial field F_0 is less than n *i e.*, $[E_0 : F_0] < n$.

Let $[E : F]=n > 1$, where E is a splitting field of $f(x)$ over F. If $f(x)$ resolves into a product of linear factors over F, then F will be a splitting field for $f(x)$ and so we cannot have $[E : F] > 1$ because $E=F \Rightarrow [E : F]=1$. Therefore $f(x)\in F[x]$ must have an irreducible factor $p(x)\in F[x]$ of degree $r > 1$. Let $p'(t)$ be the corresponding irreducible factor of $f'(t)$.

Now E is a splitting field for $f(x)\in F[x]$

$\Rightarrow$ a full complement of roots of $f(x)$ are in E

$\Rightarrow$ a full complement of roots of $p(x)$ are in E

[$\because$ $p(x)$ is a factor of $f(x)$]

$\Rightarrow$ $\exists$ $v\in E$ such that $p(v)=0$

$\Rightarrow$ $[F(v) : F]=r=\deg p(x)$ [by theorem 7].

Similarly there is a $w\in E'$ such that $p'(w)=0$.

By theorem 21 there is an isomorphism σ of $F(v)$ onto $F'(w)$ such that $\alpha\sigma=\alpha\psi=\alpha'$ for every $\alpha\in F$.

Now $[E : F]=[E : F(v)]\,[F(v) : F]$

$$\Rightarrow [E : F(v)]=\frac{[E : F]}{[F(v) : F]}=\frac{n}{r} < n \qquad [\because\ r>1]$$

Now let $F_0=F(v)$ and $F_0'=F'(w)$. Since $F_0 \supset F$, therefore $f(x)\in F[x]$ can also be regarded as $f(x)\in F_0[x]$. We claim that E is a splitting field for $f(x)\in F_0[x]$. Obviously $f(x)\in F_0[x]$ resolves into a product of linear factors over E. No proper subfield of E containing F_0 and hence F can split $f(x)$ into linear factors because we have assumed that E is a splitting field of $f(x)$ over F. Hence E is a splitting field for $f(x)\in F_0[x]$. Similarly E' is a splitting field for $f'(t)\in F_0'[t]$.

Now σ is an isomorphism of F_0 onto F_0'. E is splitting field for $f(x) \in F_0[x]$ and E' is splitting field for $f'(t) \in F_0'[t]$.

Since $[E : F_0] = [E : F(v)] < n$, therefore by our induction hypothesis there is an isomorphism ϕ of E onto E' such that

$$a\phi = a\sigma \text{ for all } a \in F_0$$

Now $\alpha \in F \Rightarrow \alpha \in F_0$. Therefore for every $\alpha \in F$, we have

$$\alpha\phi = \alpha\sigma = \alpha\psi = \alpha' \qquad [\because \ \alpha\sigma = \alpha\psi = \alpha' \ \forall \ \alpha \in F]$$

The proof is now complete by induction.

Corollary. Uniqueness of the splitting field. *Any two splitting fields of the same polynomial over a given field F are isomorphic by an isomorphism leaving every element of F fixed*

(I.A S. 1973; Meerut 76, 77, 79; Guru Nanak 82)

Proof In the proof of theorem 22 take $F' = F$ and take ψ as the identity mapping of F *i.e.*, $\alpha\psi = \alpha \ \forall \ \alpha \in F$ Then ψ is an isomorphism of F onto F and it leaves every element of F fixed Let E_1 and E_2 be two splitting fields for $f(x) \in F[x]$. In the above theorem take $E_1 = E$ and $E_2 = E'$. Then E_1 and E_2 are isomorphic by an isomorphism leaving every element of F fixed.

Solved Examples

Ex. 1. *Let F be the field of rational numbers. Determine the degree of the splitting field of the polynomial* $x^3 - 2$ *over F.*

(Meerut 1981, 84)

Solution. F is the field of rational numbers.

Let $f(x) = x^3 - 2 \in F[x]$. In the field of complex numbers the three roots of $f(x)$ are $2^{1/3}$, $\omega 2^{1/3}$, $\omega^2 2^{1/3}$, where $\omega = (-1 + i\sqrt{3})/2$ and where $2^{1/3}$ is a real cube root of 2. The polynomial $f(x)$ is irreducible over F. We have $\deg f(x) = 3$. Also $2^{1/3}$ is a root of $f(x)$. Therefore $2^{1/3}$ is algebraic over F of degree 3. Therefore

$$[F(2^{1/3}) : F] = 3.$$

Let E be the splitting field of $f(x)$ over F. The field $F(2^{1/3})$ cannot split $f(x)$ because as a subfield of the real field it cannot contain the complex, but not real, number $\omega 2^{1/3}$. Therefore $F(2^{1/3})$ will be a proper subfield of E. So we have

$$[E : F] > [F(2^{1/3}) : F] = 3.$$

By theorem 17, $[E : F] \leqslant 3\,! = 6$.

Also $[E : F] = [E : F(2^{1/3})]\ [F(2^{1/3}) : F]$

$\Rightarrow [F(2^{1/3}) : F]$ is a divisor of $[E : F] \Rightarrow 3$ is a divisor of $[E : F]$.

Now $[E : F] \leqslant 6$, $[E : F] > 3$ and 3 is a divisor of $[E : F]$.

Therefore we must have $[E : F] = 6$.

Ex. 2. *Let F be the field of rational numbers and*

$$f(x) = x^4 + x^2 + 1 \in F[x].$$

Show that $F(\omega)$ where $\omega=(-1+i\sqrt{3})/2$ is a splitting field of $f(x)$. Also determine the degree of the splitting field of $f(x)$ over F.

Solution. We have

$$x^4+x^2+1=(x^2+1)^2-x^2=(x^2+x+1)\,(x^2-x+1).$$

In some extension of F if α is a root of x^2+x+1, then $-\alpha$ is a root of x^2-x+1 in that extension. Thus if any extension of F splits x^2+x+1, then it will also split x^2-x+1 and consequently it will split x^4+x^2+1. Thus the splitting field of $x^4+x^2+1\in F[x]$ is the same as that of $x^2+x+1\in F[x]$.

Let $f(x)=x^2+x+1$. In the field of complex numbers the two roots of x^2+x+1 are ω, ω^2. Since $f(x)$ is irreducible over F and $\deg f(x)=2$, therefore in any extension of F of degree less than 2, $f(x)$ cannot have a root. So if E is the splitting field of $f(x)$, then $[E:F]\geqslant 2$. By theorem 17, $[E:F]\leqslant 2\,!=2$. So we must have $[E:F]=2$.

The field $F(\omega)$ contains a root ω of $f(x)$. Since $\omega\in F(\omega)$ $\Rightarrow \omega^2\in F(\omega)$, therefore $F(\omega)$ contains both the roots ω and ω^2 of $f(x)$. Thus $F(\omega)$ splits x^2+x+1.

The polynomial $f(x)$ is irreducible over F. We have $\deg f(x)=2$. Also ω is a root of $f(x)$. Therefore ω is algebraic over F of degree 2. So we have $[F(\omega):F]=2$.

Hence $F(\omega)$ is a splitting field of $f(x)$.

Ex. 3. *If p is a prime number prove that the splitting field over F, the field of rational numbers, of the polynomial x^p-1 is of degree $p-1$.* **(I.A.S. 1972)**

Solution. The polynomial $f(x)=x^p-1\in F[x]$ can be factored over F as

$$f(x)=(x-1)\,(x^{p-1}+x^{p-2}+\ldots+x^2+x+1).$$

Let $\quad q(x)=x^{p-1}+x^{p-2}+\ldots+x^2+x+1\in F[x]$.

Since 1 will belong to any extension of F, therefore the splitting field of $q(x)$ will also be the splitting field of $f(x)$.

Since p is prime, therefore $q(x)$ is irreducible over F. [See Ex. 3 page 393] We have $\deg q(x)=p-1$. If α is a root of $q(x)$ in some extension of F, then α is algebraic over F of degree $p-1$. So $\quad [F(\alpha):F]=p-1$.

Now $F(\alpha)$ is a minimal extension of F containing α. Therefore no extension of F of degree less than $p-1$ can have a root of $q(x)$. So if E is the splitting field of $q(x)$, then we must have

$$[E:F]\geqslant p-1.$$

We claim that $[E : F]=p-1$.

Let us solve $x^p-1=0$ in **C**, the field of complex numbers.

We have $x^p=1$

$\Rightarrow x=(1+0i)^{1/p}=(\cos 0+i \sin 0)^{1/p}$

$=(\cos 2n\pi+i \sin 2n\pi)^{1/p}$, where n is any integer

$$=\cos \frac{2n\pi}{p}+i \sin \frac{2n\pi}{p}=e^{i2n\pi/p}.$$

Putting $n=0, 1, \ldots, p-1$, we get

$$1, e^{2\pi i/p}, e^{4\pi i/p}, \ldots, e^{[2\pi(p-1)i]/p}$$

as the p roots of x^p-1 in **C**.

These p roots form a cyclic multiplicative group of prime order p.

Let $\beta=e^{2\pi i/p}$. Then the elements of this group can be written as

$$1, \beta, \beta^2, \ldots, \beta^{p-1}.$$

The field $F(\beta)$ contains β. Also $1, \beta^2, \ldots, \beta^{p-1} \in F(\beta)$. Thus the field $F(\beta)$ splits x^p-1.

Since β is a root of x^p-1 and $\beta \neq 1$, therefore β is also a root of $x^{p-1}+\ldots+x^2+x+1$ which is irreducible over F and is of degree $p-1$.

So $[F(\beta) : F]=p-1$.

Hence $F(\beta)$ is the splitting field of x^p-1.

Ex. 4 *Let F be the field of rational numbers. Determine the degree of the splitting field of the polynomial* x^5-1 *over F.*

Solution. Since 5 is prime, therefore proceed as in Ex. 3.

Ex. 5. *If E is an extension of F and if* $f(x) \in F[x]$ *and if* ϕ *is an automorphism of E leaving every element of F fixed, prove that* ϕ *must take a root of* $f(x)$ *lying in E into a root of* $f(x)$ *in E.*

(Meerut 1979)

Solution. Let $f(x)=a_0+a_1x+\ldots+a_nx^n$, where $a_0, a_1, \ldots, a_n \in F$.

E is an extension of F and ϕ is an automorphism of E leaving every element of F fixed. Thus $\phi(a)=a \ \forall \ a \in F$.

Let α be a root of $f(x)$ in E. Then $a_0+a_1\alpha+\ldots+a_n\alpha^n=0$.

To prove that $\phi(\alpha)$ is also a root of $f(x)$ in E. Obviously $\phi(\alpha) \in E$ because ϕ is a mapping from E onto E. Let $\phi(\alpha)=\beta$. If r is any positive integer, we have

$\phi(\alpha^r)=\phi(\alpha\ \alpha\ \alpha \ldots r$ times$)$

$=\phi(\alpha)\ \phi(\alpha) \ldots r$ times [$\because$ ϕ is an automorphism]

$=\beta\,\beta\,\beta\ldots r$ times $=\beta^r$.

Also ϕ will map 0 onto 0 because it is an automorphism. Thus $\phi(0)=0$.

Now $a_0+a_1\alpha+\ldots+a_n\alpha^n=0$

$\Rightarrow \phi(a_0+a_1\alpha+\ldots+a_n\alpha^n)=\phi(0)$

$\Rightarrow \phi(a_0)+\phi(a_1)\,\phi(\alpha)+\ldots+\phi(a_n)\,\phi(\alpha^n)=0$

[Since ϕ is an automorphism, therefore it preserves additions and multiplications. Also $\phi(0)=0$]

$\Rightarrow a_0+a_1\beta+a_2\beta^2+\ldots+a_n\beta^n=0$ [since $\phi(\alpha^r)=\beta^r$ and $a_0, a_1, \ldots, a_n \in F$. Note $\phi(a)=a \ \forall\ a \in F$]

$\Rightarrow \beta$ is a root of $f(x)$.

Ex. 6. *Using the result of Ex. 5, prove that if the complex number z is a root of the polynomial $p(x)$ having real coefficients then $\bar{z}$, the complex conjugate of z, is also a root of $p(x)$.*

Solution. Let **R** be the field of real numbers and **C** be the field of complex numbers. Then **C** is an extension of **R**. We have

$$\mathbf{C}=\{a+ib : a, b \in \mathbf{R}\}.$$

Let $p(x)\in \mathbf{R}[x]$ and let $z=x+iy$ be a root of $p(x)$ in **C**.

Then to prove that $\bar{z}=x-iy$ is also a root of $p(x)$ in **C**.

Let ϕ be a mapping from **C** into **C** defined as

$$\phi(x+iy)=x-iy \ \forall\ x, y \in \mathbf{R}.$$

We claim that ϕ is an automorphism of **C**.

ϕ is one-one.

We have $\phi(a+ib)=\phi(c+id)$ where $a, b, c, d \in \mathbf{R}$

$\Rightarrow a-ib=c-id \Rightarrow a=c,\ b=d \Rightarrow a+ib=c+id.$

ϕ is onto. Let $a+ib$ be any element of **C**. Then $a-ib\in \mathbf{C}$, and we have $\phi(a-ib)=a+ib$.

ϕ preserves additions and multiplications.

We have $\phi[(a+ib)+(c+id)]=\phi[(a+c)+i(b+d)]$

$=(a+c)-i(b+d)=(a-ib)+(c-id)=\phi(a+ib)+\phi(c+id)$.

Also $\phi[(a+ib)(c+id)]=\phi[(ac-bd)+i(ad+bc)]$

$=(ac-bd)-i(ad+bc)=(a-ib)(c-id)=[\phi(a+ib)][\phi(c+id)]$.

Hence ϕ is an automorphism of **C**.

If $a\in \mathbf{R}$, then as an element of **C** it can be written as $a+i0$. We have $\phi(a)=\phi(a+i0)=a-i0=a$.

Thus ϕ leaves every element of **R** fixed

Hence by the result of Ex. 5 if z is a root of $p(x)$ in **C**, then $\phi(z)=\bar{z}$ is also a root of $p(x)$ in **C**.

Ex. 7. *Using the result of Ex. 5 prove that if m is an integer which is not a perfect square and if $\alpha+\beta\sqrt{m}$ (α, β rational) is the* *root of a polynomial $p(x)$ having rational coefficients, then $\alpha-\beta\sqrt{m}$ is also a root of $p(x)$.*

Solution. Let Q be the field of rational numbers and let

$$K=\{x+y\sqrt{m} : x, y \in Q\}.$$

It can be easily seen that $(K, +, \cdot)$ is a field. Also K is an extension of Q. Let $p(x) \in Q[x]$ and let $\alpha+\beta\sqrt{m}$ be a root of $p(x)$ in K. Then to prove that $\alpha-\beta\sqrt{m}$ is also a root of $p(x)$ in K.

Let ϕ be a mapping from K into K defined as

$$\phi(x+y\sqrt{m})=x-y\sqrt{m} \quad \forall \; x, y \in Q.$$

We claim that ϕ is an automorphism of K.

ϕ is one-one.

We have $\phi(a+b\sqrt{m})=\phi(c+d\sqrt{m})$, $a, b, c, d \in Q$

$\Rightarrow a-b\sqrt{m}=c-d\sqrt{m} \Rightarrow a=c, b=d \Rightarrow a+b\sqrt{m}=c+d\sqrt{m}.$

ϕ is onto.

Let $a+b\sqrt{m}$ be any element of K. Then $a-b\sqrt{m} \in K$. We have $\phi(a-b\sqrt{m})=a+b\sqrt{m} \Rightarrow \phi$ is onto.

ϕ preserves additions and multiplications. We have

$$\begin{aligned}\phi[(a+b\sqrt{m})+(c+d\sqrt{m})]&=\phi\,[(a+c)+(b+d)\sqrt{m}]\\ &=(a+c)-(b+d)\sqrt{m}=(a-b\sqrt{m})+(c-d\sqrt{m})\\ &=\phi\,(a+b\sqrt{m})+\phi(c+d\sqrt{m}).\end{aligned}$$

Also

$$\begin{aligned}\phi\,[(a+b\sqrt{m})\,(c+d\sqrt{m})]&=\phi[(ac+bdm)+(ad+bc)\sqrt{m}]\\ &=(ac+bdm)-(ad+bc)\sqrt{m}=(a-b\sqrt{m})\,(c-d\sqrt{m})\\ &=[\phi(a+b\sqrt{m})]\;[\phi(c+d\sqrt{m})].\end{aligned}$$

Hence ϕ is an automorphism of K. If $a \in Q$, then as an element of K it can be written as $a+0\sqrt{m}$. We have

$$\phi(a)=\phi(a+0\sqrt{m})=a-0\sqrt{m}=a.$$

Thus ϕ leaves every element of Q fixed.

Hence by the result of Ex. 5 if $\alpha+\beta\sqrt{m}$ is a root of $p(x)$ in K, then $\phi\,(\alpha+\beta\sqrt{m})=\alpha-\beta\sqrt{m}$ is also a root of $p(x)$ in K.

Ex. 8. *Show that the polynomials x^2+3 and x^2+x+1 have same splitting field over* Q, *the field of rational numbers.* **(Meerut 1974)**

§ 6. More about roots

Derivative of a polynomial over a field. Definition.

Let $f(x)=a_0+a_1x+a_2x^2+\ldots+a_{n-1}x^{n-1}+a_nx^n$ be a polynomial over a field F. Then the derivative of $f(x)$ denoted by $f'(x)$, is defined as the polynomial

$$f'(x)=a_1+2a_2x+\ldots+(n-1)\,a_{n-1}x^{n-2}+na_nx^{n-1} \text{ in } F[x].$$

Note. Here by 2 we mean $1+1$, by n we mean $1+1+\ldots$upto n times. The field F is arbitrary. It may be a field of finite characteristic or it may be a field of characteristic zero or infinite. If F is a field of finite characteristic p, then the derivative of the polynomial x^p is $px^{p-1}=0\,x^{p-1}=0$. Recall that $p=1+1+1+\dots$ upto p times $=0$ because the characteristic of the field is p. Thus we see that even the derivative of a non-constant polynomial can be zero if field F is a field of finite characteristic.

Theorem 23. *Let F be a field and let $f(x)\in F[x]$ be such that $f'(x)=0$. Prove that*

(i) If charecteristic $F=0$, then $f(x)=a\in F$ i.e., $f(x)$ is a constant polynomial.

(ii) If characteristic $F=p\neq 0$, then $f(x)=g(x^p)$ for some polynomial $g(x)\in F[x]$ i.e., $f(x)$ is a polynomial in x^p.

(Kanpur 1970)

Proof. Let $f(x)=a_0+a_1x+a_2x^2+\dots+a_ix^i+\dots+a_nx^n$ be a polynomial over a field F.

We have $f'(x)=a_1+2a_2x+\dots+ia_ix^{i-1}+\dots+na_nx^{n-1}$.

It is given that $f'(x)=0=$ zero polynomial. Therefore

$$a_1+2a_2x+\dots+ia_ix^{i-1}+\dots+na_nx^{n-1}=0+0x+\dots+0x^{n-1}.$$

By definition of equality of polynomials, we get

$$a_1=0,\ 2a_2=0,\dots,\ ia_i=0,\dots,\ na_n=0.$$

Case (i). Characteristic $F=0$.

We have $ia_i=0$

$\Rightarrow a_i+a_i+\dots$ upto i times $=0$

$\Rightarrow$ order of a_i regarded as an element of the additive group of F is $\leqslant i$

$\Rightarrow$ order of a_i is finite

$\Rightarrow a_i=0$, because in a field of characteristic 0 each non-zero element is of infinite order and zero is the only element of finite order 1.

Thus $a_1=0,\ a_2=0,\dots,\ a_i=0,\dots,\ a_n=0$.

$\therefore\ f(x)=a_0\in F\Rightarrow f(x)$ is contstant.

Case (ii). Characteristic $F=p\neq 0$.

In this case each non-zero element of F is of order p and the ero element of F is of order 1.

From the theory of groups we know that if a is an element of rder p, then $a^m=e\Rightarrow p$ is a divisor of m.

Here $e=0$, the identity of the additive group of F.

Now $\quad ia_i=0$

$\Rightarrow$ either $a_i=0$ or if $a_i\neq 0$ then p should be a divisor of i

$\Rightarrow$ either $a_i=0$ or if $a_i\neq 0$ then $i=\lambda p$ where λ is some positive integer.

Therefore in this case the term x^i will occur in $f(x)$ only if it is of the form $(x^p)^\lambda$. Thus $f(x)$ will be a polynomial in x^p.

Theorem 24. *For any* $f(x)$, $g(x) \in F[x]$ *and any* $\alpha \in F$

(*i*) $[f(x)+g(x)]' = f'(x)+g'(x)$

(*ii*) $[\alpha f(x)]' = \alpha f'(x)$

(*iii*) $[f(x)\, g(x)]' = f'(x)\, g(x)+f(x)\, g'(x)$.

Proof. (i) Let $f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n$

and $g(x)=b_0+b_1x+b_2x^2+\ldots+b_mx^m$.

Without loss of generality, we can take $n \geqslant m$.

We have $f(x)+g(x)=(a_0+b_0)+(a_1+b_1)\,x+(a_2+b_2)\,x^2+\ldots$
$+(a_m+b_m)\,x^m+(a_{m+1}+0)\,x^{m+1}+\ldots+(a_n+0)\,x^n$, where $a_{m+1}, \ldots, a_n$ are all zero if $n=m$.

Now $[f(x)+g(x)]'=(a_1+b_1)+2\,(a_2+b_2)\,x+\ldots$
$\ldots+m(a_m+b_m)\,x^{m-1}+(m+1)\,a_{m+1}\,x^m+\ldots+na_nx^{n-1}$.

Now if $a, b \in F$ and k is any integer, then $k\,(a+b)=ka+kb$.

$\therefore$ $[f(x)+g(x)]'$

$=[a_1+2a_2x+\ldots+ma_mx^{m-1}+(m+1)\,a_{m+1}\,x^m+\ldots+na_n\,x^{n-1}]$
$+[b_1+2b_2x+\ldots+mb_mx^{m-1}]$

$=f'(x)+g'(x)$.

(ii) Let $f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n$.

We have $\alpha f(x)=\alpha a_0+\alpha a_1x+\alpha a_2x^2+\ldots+\alpha a_nx^n$.

$\therefore$ $[\alpha f(x)]'=\alpha a_1+2\,(\alpha a_2)\,x+\ldots+n\,(\alpha a_n)\,x^{n-1}$
$=\alpha[a_1+2a_2x+\ldots+na_nx^{n-1}]=\alpha f'(x)$.

(iii) Since the product of $f(x)$ with $g(x)$ is a linear combination over F of powers of x, therefore by virtue of results (i) and (ii) it is sufficient to prove the result (iii) for products of two powers of x. Let $f(x)=x^r$, $g(x)=x^s$ where r and s are positive integers.

We have $f(x)\, g(x)=x^r\, x^s=x^{r+s}$. Therefore

$[f(x)\, g(x)]'=(r+s)\, x^{r+s-1}=rx^{r+s-1}+sx^{r+s-1}$
$=(r\, x^{r-1})\, x^s+x^r\,(sx^{s-1})=f'(x)\, g(x)+f(x)g'(x)$.

Condition for multiple roots.

Theorem 25. *The polynomial* $f(x) \in F[x]$ *has a multiple root if and only if* $f(x)$ *and* $f'(x)$ *have a nontrivial (that is, of positive degree) common factor.*

Proof. Before proving the main theorem let us first prove an important result namely :

If $f(x)$ *and* $g(x)$ *in* $F[x]$ *have a non-trivial common factor in* $K[x]$ *where* K *is some extension of* F, *then they must have a non-trivial common factor in* $F[x]$.

We shall prove this result by contradiction. Suppose $f(x)$ and $g(x)$ have a non-trivial common factor in $K[x]$ but they have no non-trivial common factor in $F[x]$. Then $f(x)$ and $g(x)$ are relatively prime as elements in $F[x]$. Therefore there exist two polynomials $a(x)$ and $b(x)$ in $F[x]$ such that

$$a(x)\,f(x)+b(x)\,g(x)=1.$$

Since $K \supseteq F$, therefore $a(x)$, $b(x)$, $f(x)$, $g(x)$ can all be taken as polynomials in $K[x]$.

Therefore $a(x)\,f(x)+b(x)\,g(x)=1$ implies that $f(x)$ and $g(x)$ are relatively prime as elements in $K[x]$. Note that if $f(x)$ and $g(x)$ have a common factor of positive degree then this factor must divide 1 which is impossible. Thus we get a contradiction. Hence $f(x)$ and $g(x)$ must have a nontrivial common factor in $F[x]$.

Now we come to the proof of the main theorem. By virtue of the result we just proved we can assume without loss of generality that the roots of $f(x)$ all lie in F (otherwise extend F to K the splitting field of $f(x)$).

Now suppose that α is a multiple root of $f(x)$ of multiplicity $m > 1$. Then $(x-\alpha)^m$ is a divisor of $f(x)$. Therefore

$f(x)=(x-\alpha)^m\,q(x)$. We have

$$f'(x)=[(x-\alpha)^m]'\,q(x)+(x-\alpha)^m\,q'(x)$$
$$=m\,(x-\alpha)^{m-1}\,q(x)+(x-\alpha)^m\,q'(x)$$

[Note that if $m > 1$, then the derivative of $(x-\alpha)^m$ is $m\,(x-\alpha)^{m-1}$]

$$=(x-\alpha)\,r(x), \text{ since } m > 1.$$

Thus $(x-\alpha)$ is a common factor of $f(x)$ and $f'(x)$. So $f(x)$ and $f'(x)$ have a non-trivial common factor.

Converse. Suppose that $f(x)$ and $f'(x)$ have a non-trivial common factor. Then to prove that $f(x)$ has a multiple root.

Suppose that $f(x)$ has no multiple root.

Let the roots of $f(x)$ be $\alpha_1, \alpha_2, \dots, \alpha_n$ where the α_i's are all distinct. Without loss of generality we can assume $f(x)$ to be monic. Then $f(x)=(x-\alpha_1)(x-\alpha_2)\dots(x-\alpha_n)$. Therefore

$$f'(x)=\sum_{i=1}^{n}(x-\alpha_1)\dots\overline{(x-\alpha_i)}\dots(x-\alpha_n)$$

where the bar —— denotes the term is omitted.

Our claim is that no root of $f(x)$ is a root $f'(x)$, for

$$f'(\alpha_i)=\prod_{j\neq i}(\alpha_i-\alpha_j)\neq 0$$

since $\alpha_1, \dots, \alpha_n$ are all distinct. Now if $f(x)$ and $f'(x)$ have a non-trivial common factor, then they have a common root which is any root of their common factor. Therefore $f(x)$ and $f'(x)$ have no common root implies that they have no nontrivial common factor. Thus we have proved that if $f(x)$ has no multiple root, then $f(x)$ and $f'(x)$ have no non-trivial common factor.

Hence if $f(x)$ and $f'(x)$ have a nontrivial common factor, then $f(x)$ must have a multiple root. This proves the other side of the theorem.

Theorem 26. *If $f(x)\in F[x]$ is irreducible, then :*

(i) If the characteristic of F is 0, $f(x)$ has no multiple roots.

(ii) If the characteristic of F is $p\neq 0$, $f(x)$ has a multiple root only if it is of the form $f(x)=g(x^p)$ for some $g(x)\in F[x]$.

Proof. Suppose $f(x)$ has a multiple root. Then by theorem 25, $f(x)$ and $f'(x)$ have a non-trivial common factor in $F[x]$. Since $f(x)$ is irreducible in $F[x]$, therefore its only non-trivial factor in $F[x]$ is $f(x)$. Thus $f(x)$ is a factor of $f'(x)$. Now if $f'(x)\neq 0$, then deg $f'(x)$ is less than that of $f(x)$ and so $f(x)$ cannot be a factor of $f'(x)$. Therefore $f(x) \mid f'(x)$ is possible only if $f'(x)=0$. In characteristic 0 this implies that $f(x)$ is a constant, which has no roots; in characteristic $p\neq 0$, this forces $f(x)=g(x^p)$ for some $g(x)\in F[x]$.

Theorem 27. *If F is a field of characteristic $p\neq 0$, then the polynomial*

$$x^{p^n}-x\in F[x],$$

for $n\geqslant 1$, has distinct roots.

Proof. Let $f(x)=x^{p^n}-x$. Then $f'(x)=p^n x^{p^n-1}-1$.

Now by $p\in F$, we mean $1+1+1+$ ...upto p times. Since F is of characteristic p, therefore the order of 1 as an element of the additive group of F is p. Therefore $p=1+1+$ ...upto p times $=0$.

Hence $p^n=0$. Therefore $f'(x)=-1$.

Now we see that $f(x)$ and $f'(x)$ have no nontrivial common factor. Therefore by theorem 25, $f(x)$ has no multiple roots.

Theorem 28. *If F is of characteristic 0 and if a, b are algebraic over F, then there exists an element $c\in F(a, b)$ such that*

$F(a, b)=F(c)$. **(Meerut 1973, 78)**

Or

If F is of characteristic 0 and if a, b are algebraic over F, then $F(a, b)$ is a simple extension of F.

Proof. It is given that the elements a and b are algebraic over F. Let $f(x)$ and $g(x)$ be the irreducible polynomials over F

satisfied by a and b respectively. Let $\deg f(x)=m$ and $\deg g(x)=n$. Suppose E is an extension of F in which both $f(x)$ and $g(x)$ split completely. Since $f(x)$ is irreducible and characteristic $F=0$, therefore by theorem 26 all the roots of $f(x)$ are distinct. Similarly all the roots of $g(x)$ are distinct. Let $a_1, a_2, .., a_m$ be the roots of $f(x)$ and $b_1, b_2, ..., b_n$ be the roots of $g(x)$. For the sake of convenience let us put $a_1=a$, $b_1=b$.

If $j\neq 1$, then $b_j\neq b_1=b$. Therefore if in E, we have

$$a_i+\lambda b_j=a+\lambda b,\ i=1,...m,\ j=2,...,n,$$

then $\lambda=\dfrac{a_i-a}{b-b_j}$ which is a unique element of E.

Thus for each pair of values of i and j provided $j\neq 1$, the equation $a_i+\lambda b_j=a+\lambda b$ has only one solution in E. By putting $i=1,..., m$, $j=2, . . , n$, the equation $a_i+\lambda b_j=a+\lambda b$ can take only a finite number of forms. Thus there are only a finite number of elements of E such that

$$a_i+\lambda b_j=a+\lambda b, \text{ where } i=1,..., m, j=2,..., n.$$

Now F has an infinite number of elements since characteristic $F=0$. Each element of F is in E. Therefore we must have an element $\gamma\in F$ such that

$$a_i+\gamma b_j\neq a+\gamma b \text{ for all } i \text{ and for all } j\neq 1.$$

Let $c=a+\gamma b$. Since $a, b, \gamma\in F(a, b)$, therefore $c\in F(a, b)$. We shall prove that $F(a, b)=F(c)$.

Since $c\in F(a, b)$ therefore $F(c)\subseteq F(a, b)$.

Now it remains to prove that $F(a, b)\subseteq F(c)$.

For this we shall prove that both a and b are in $F(c)$.

Now b satisfies the polynomial $g(x)$ over F. The polynomial $g(x)$ over F can also be regarded as a polynomial over $F(c)$ because $F\subseteq F(c)$. Let us put $F(c)=K$. Let $h(x)=f(c-\gamma x)$. Since $c\in K$, and $\gamma\in F\Rightarrow\gamma\in K$, therefore $h(x)$ is a polynomial in $K[x]$. We have

$$\begin{aligned} h(b)&=f(c-\gamma b)\\ &=f(a) && [\because\ c=a+\gamma b]\\ &=0, \text{ since } a \text{ is a root of } f(x). \end{aligned}$$

Thus b satisfies both the polynomials $g(x)$ and $h(x)$ in $K[x]$. Therefore $x-b$ is a common factor of $g(x)$ and $h(x)$ in some extension E of K. [Note that $K=F(c)\subseteq F(a, b)\subseteq E$ Hence E is an extension of K]. We shall prove that $x-b$ is the greatest common divisor of $g(x)$ and $h(x)$ in $K[x]$. Let $b_j\neq b$ be another root of $g(x)$. Then

$h(b_j)=f(c-\gamma b_j)$

$\neq 0$, since by our choice of γ, $c-\gamma b_j$ for $j\neq 1$ avoids all roots a_i of $f(x)$.

[Note that $a_i+\gamma b_j\neq a+\gamma b=c$ for all i and for all $j\neq 1$].

Therefore b_j is not a root of $h(x)$. Thus any factor of $g(x)$ in $E[x]$ other than $x-b$ is not a factor of $h(x)$.

Also $g(x)$ has all its roots distinct. So $(x-b)^r$, where $r\geqslant 2$, is not a divisor of $g(x)$. Hence $x-b$ is the greatest common divisor of $g(x)$ and $h(x)$ in some extension E of K. Since $g(x)$ and $h(x)$ have a nontrivial common factor over some extension of K, therefore they must have a nontrivial common factor over K. [See initial remark in the proof of theorem 25]. Consequently they must have a nontrivial greatest common divisor over K which must be a divisor of $x-b$. But the degree of $x-b$ is 1. Therefore $x-b$ itself is the greatest common divisor of $g(x)$ and $h(x)$ in $K[x]$.

Hence $b\in K=F(c)$.

Now $c,\ \gamma,\ b\in F(c)\Rightarrow c-\gamma b\in F(c)\Rightarrow a\in F(c)$.

Thus both a, b are in $F(c)$. We have

$$a,\ b\in F(c)\Rightarrow F(a,b)\subseteq F(c).$$

Finally $F(a,b)\subseteq F(c),\ F(c)\subseteq F(a,b)\Rightarrow F(a,b)=F(c)$.

This proves the theorem.

Theorem 29. *Any finite extension of a field of characteristic 0 is a simple extension.*

Proof. Let K be a finite extension of a field of characteristic 0. Then K is an algebraic extension of F and can be obtained by adjoining a finite number of algebraic elements of F. Let

$$K=F(\alpha_1,\ \alpha_2,\ldots,\ \alpha_n).$$

We are to prove that there is an element $c\in K$ such that $K=F(c)$.

$$\begin{aligned}\text{We have } K&=F(\alpha_1,\ \alpha_2,\ldots,\ \alpha_{n-2},\ \alpha_{n-1},\ \alpha_n)\\ &=(F(\alpha_1,\ \alpha_2,\ldots,\ \alpha_{n-2}))\ (\alpha_{n-1},\ \alpha_n)\\ &=(F(\alpha_1,\ \alpha_2,\ldots,\ \alpha_{n-2}))\ (d)\ \text{where}\end{aligned}$$

$$d\in(F(\alpha_1,\ \alpha_2,\ldots,\ \alpha_{n-2}))\ (\alpha_{n-1},\ \alpha_n)=F(\alpha_1,\ \alpha_2,\ldots,\ \alpha_{n-1},\ \alpha_n)=K.$$

[We have applied theorem 28 by taking

$F=F(\alpha_1,\ ..,\ \alpha_{n-2})$ and $\alpha_{n-1}=a,\ \alpha_n=b$]

$$=F(\alpha_1,\ \alpha_2,\ldots,\ \alpha_{n-2},\ d).$$

By repeated application of theorem 28, in the manner shown above, we shall get after a finite number of steps

$$K = F(c) \text{ where } c \in F(c) = K.$$

Hence K is a simple extension of F.

Separable element. Definition. *An element a in an extension K of F is called separable over F if it satisfies a polynomial over F having no multiple roots.*

Separable extension. Definition. *An extension K of a field F is called separable over F if each element of K is separable over F.*

(Meerut 1981, 82, 83)

Perfect field. Definition. *A field F is called perfect if all finite extensions of F are separable.*

Theorem 30. *Show that any field of characteristic* 0 *is perfect.*

Proof. Let F be a field of characteristic 0. The field F will be a perfect field if every finite extension of F is separable over F. Let K be any finite extension of F. Then to prove that K is separable over F. Let $a \in K$ be arbitrary. Then K will be separable over F if a is separable over F *i.e.*, if a satisfies a polynomial over F having no multiple roots.

Let $$[K : F] = n.$$

Then a satisfies a non-zero polynomial in $F[x]$ of degree at most n. [See theorem 7].

Suppose a satisfies $f(x) \in F[x]$. Then $f(a) = 0$. By the unique factorization theorem for polynomials over a field the polynomial $f(x) \in F[x]$ can be expressed as

$$f(x) = q_1(x)\, q_2(x) \dots q_m(x)$$

where $q_1(x), \dots, q_m(x)$ are irreducible elements in $F[x]$. We have

$$\begin{aligned} & f(a) = 0 \\ \Rightarrow\ & q_1(a)\, q_2(a) \dots q_m(a) = 0 \\ \Rightarrow\ & q_i(a) = 0 \text{ for at least one } i \text{ between } 1 \text{ and } m \\ \Rightarrow\ & a \text{ satisfies an irreducible polynomial } q_i(x) \text{ in } F[x]. \end{aligned}$$

Since characteristic F is 0 and $q_i(x) \in F[x]$ is irreducible, therefore $q_i(x)$ cannot have multiple roots. [See theorem 26]

Thus a satisfies a polynomial over F having no multiple roots.

Hence F is a perfect field.

§ 7. The Elements of Galois Theory.

Let $p(x)$ be a polynomial in $F[x]$ where F is any field. In this section we shall associate with $p(x)$ a group called the **Galois group** of $p(x)$. This Galois group will turn out to be a certain permutation group of the roots of the polynomial.

In this section we shall assume that all our fields are of characteristic 0. Then we can make free use of theorems 28 and

29. However most of the results will be true for an arbitrary field *i.e.*, they will also be true for fields having finite characteristic. Therefore we shall prefer to mention in the statement of the theorem whether the field is of characteristic 0 or an arbitrary field. If we write the words 'any field' then the result will also be true for finite characteristic fields.

Automorphism of a field. Definition.

A one-one mapping σ of a field K onto itself is called an automorphism of K if

$\sigma(a+b)=\sigma(a)+\sigma(b)$ *and* $\sigma(ab)=\sigma(a)\,\sigma(b)$ *for all* $a, b\in K$.

Two automorphisms σ_1 and σ_2 of K are said to be equal if

$$\sigma_1(a)=\sigma_2(a)\ \forall\ a\in K.$$

Thus σ_1 and σ_2 will be distinct if $\sigma_1(a)\neq\sigma_2(a)$ for some element a in K.

Note. If σ is an automorphism of a field K, then

(i) $\sigma(0)=0$, (ii) $\sigma(-a)=-\sigma(a)$,

(iii) $\sigma(1)=1$, (iv) $\sigma(-1)=-1$,

(v) If $a\neq0$, then $\sigma(a^{-1})=[\sigma\ a)]^{-1}$.

To prove these results see the corresponding results on isomorphism of rings. Automorphism is nothing but a special type of isomorphism.

Group of automorphisms of a field.

Theorem 31. *Let A (K) be the collection of all automorphisms of a field K. Then it can be easily seen that A (K) is a group with respect to the operation known as composite of two functions.*

Proof. Let $A(K)$ be the collection of all automorphisms of a field K. Then $A\ (K)=\{f : f$ is an automorphism of $K\}$.

We shall prove that $A(K)$ is a group with respect to composite of mappings as the operation.

Closure property. Let $f, g\in A(K)$. Then f, g are one-one mappings of K onto itself. Therefore $g\,f$ is also a one-one mapping of K onto itself.

If a, b be any two elements of K, we have

$$(g\,f)\ (ab)=g[f(ab)]=g\ [f(a)\,f(b)]=g\ [f(a)]\ g\ [f\,(b)]$$
$$=[(g\,f)\ (a)]\ [(g\,f)\ (b)].$$

Also $(gf)\ (a+b)=g\ [f(a+b)]=g\ [f\,(a)+f(b)]$

$$=g\ [f(a)]+g\ [f(b)]=(g\ f)\ (a)+(gf)\ (b).$$

$\therefore$ gf is also an automorphism of K. Thus $A\ (K)$ is closed with respect to composite composition.

Associativity. We know that composite of arbitrary mappings is associative. Therefore composite of automorphisms is also associative.

Existence of identity. The identity function I on K is also an automorphism of K. Obviously I is one-one onto and if $a, b \in K$, then $I(ab) = ab = I(a)\, I(b)$, and $I(a+b) = a+b = I(a) + I(b)$.

Thus $I \in A(K)$ and if $f \in A(K)$, we have $If = f = fI$.

Existence of inverse. Let $f \in A(K)$. Since f is a one-one mapping of K onto itself, therefore f^{-1} exists and is also a one-one mapping of K onto itself. We shall show that f^{-1} is also an automorphism of K.

Let $a, b \in K$. Then $\exists\ a', b' \in K$ such that

$f^{-1}(a) = a' \Leftrightarrow f(a') = a$ and $f^{-1}(b) = b' \Leftrightarrow f(b') = b$.

We have $f^{-1}(ab) = f^{-1}[f(a')\, f(b')] = f^{-1}[f(a'b')] = a'b'$
$= f^{-1}(a)\, f^{-1}(b)$.

Also $f^{-1}(a+b) = f^{-1}[f(a') + f(b')] = f^{-1}[f(a'+b')] = a' + b'$
$= f^{-1}(a) + f^{-1}(b)$.

$\therefore$ f^{-1} is an automorphism of K and thus $f \in A(K) \Rightarrow f^{-1} \in A(K)$. Therefore each element of $A(K)$ possesses inverse.

Hence $A(K)$ is a group with respect to composite composition.

Fixed field. Definition. *Let G be a subgroup (or simply a subset) of the group of all automorphisms of a field K. Then the fixed field of G is the set of all elements $a \in K$ such that*

$$\sigma(a) = a \text{ for all } \sigma \in G.$$

The above definition of fixed field will be sensible only if we are able to show that all elements $a \in K$ such that $\sigma(a) = a$ for all $\sigma \in G$ actually form a subfield of K. We have the following theorem :

Theorem 32. *Let G be a subgroup of the group of all automorphisms of a field K. Then the fixed field of G is a subfield of K.*
(Meerut 1970; Banaras 70)

Proof. Let G be a subgroup (or simply a subset) of the group of all automorphisms of a field K. Let

$$H = \{a \in K : \sigma(a) = a\ \forall\ \sigma \in G\}.$$

To prove that H is a subfield of K.

Obviously H is not empty because at least 0 and 1 are in H. We have $\sigma(0) = 0$ and $\sigma(1) = 1$ for all $\sigma \in G$.

Let $a, b \in H$. Then $\sigma(a) = a$, $\sigma(b) = b$, for all $\sigma \in G$. For any $\sigma \in G$, we have $\sigma(a-b) = \sigma(a) - \sigma(b) = a - b$.

Thus $a, b \in H \Rightarrow a - b \in H$.

Further let $a \in H$, $0 \neq b \in H$. Then $\sigma(a) = a$, $\sigma(b) = b$ for all $\sigma \in G$. Also $0 \neq b \Rightarrow b^{-1}$ exists. For any $\sigma \in G$, we have

$\sigma(ab^{-1}) = \sigma(a)\sigma(b^{-1})$

$= [\sigma(a)]\,[\sigma(b)]^{-1}$ [Note that under an isomorphism $\sigma(b^{-1}) = [\sigma(b)]^{-1}$]

$= ab^{-1}$.

Thus $a \in H$, $0 \neq b \in H \Rightarrow ab^{-1} \in H$. Hence H is a subfield of K.

Group of automorphisms of a field K relative to a subfield F of K.

Theorem 33. *Let F be a subfield of the field K Let $G(K, F)$ be the set of those automorphisms of K which leave every element of F fixed. In other words the automorphism f of K is in $G(K, F)$ if and only if $f(\alpha) = \alpha$ for every $\alpha \in F$. Prove that $G(K, F)$ is a subgroup of the group of all automorphisms of K.*

(Banaras 1970; Meerut 77)

Proof. First we see that $G(K, F)$ is not empty because at least the identity automorphism I of K is in $G(K, F)$. Since I is the identity mapping of K, therefore $I(\alpha) = \alpha$ for every $\alpha \in F$.

Now let $f, g \in G(K, F)$. Then

$f(\alpha) = \alpha$ for every $\alpha \in F$ and $g(\alpha) = \alpha$ for every $\alpha \in F$.

If $\alpha \in F$ be arbitrary, we have

$(gf^{-1})(\alpha) = g[f^{-1}(\alpha)]$

$= g(\alpha)$ [$\because$ $f(\alpha) = \alpha \Rightarrow f^{-1}(\alpha) = \alpha$]

$= \alpha$.

Thus $gf^{-1} \in G(K, F)$.

Since $g, f \in G(K, F) \Rightarrow gf^{-1} \in G(K, F)$, therefore $G(K, F)$ is a subgroup of the group of all automorphisms of K.

Definition. *Let F be a subfield of the field K. Then the group of automorphisms of K relative to F, written as $G(K, F)$, is the set of those automorphisms of K which leave every element of F fixed.*

Note 1. $G(K, F)$ is a subgroup of the group of all automorphisms of K. If $\alpha \in K$ be such that $\alpha \in F$, then $f(\alpha) = \alpha$ for all $f \in G(K, F)$. Therefore the fixed field of $G(K, F)$ must contain F.

Note 2. Let K be a field containing the field of rational numbers F_0. Let G be the group of all automorphisms of K. Then the fixed field of G is a subfield of K. Now every subfield of K must contain F_0 because F_0 is the smallest subfield of K. Therefore the fixed field of G must contain F_0. From this we conclude that if $\alpha \in F_0$, then $\sigma(\alpha) = \alpha$ for all $\sigma \in G$. Thus every rational number is left fixed by every automorphism of K.

Theorem 34. *If K is a field and if $\sigma_1, \sigma_2, \ldots, \sigma_n$ are distinct automorphisms of K then it is impossible to find elements $a_1, a_2, \ldots, a_n$ not all 0, in K such that*

$$a_1\sigma_1(u) + a_2\sigma_2(u) + \ldots + a_n\sigma_n(u) = 0 \text{ for all } u \in K.$$

(Meerut 1974, 81, 83)

Proof. It is given that $\sigma_1, \sigma_2, \ldots, \sigma_n$ are distinct automorphisms of K. Suppose we can find elements $a_1, a_2, \ldots, a_n$, not all 0, in K such that

$$a_1\sigma_1(u) + a_2\sigma_2(u) + \ldots + a_n\sigma_n(u) = 0 \ \forall\ u \in K. \qquad \ldots(1)$$

Among all relations of the type (1) we can find a relation which has as few non-zero terms as possible. After renumbering let this minimal relation be

$$a_1\sigma_1(u) + a_2\sigma_2(u) + \ldots + a_m\sigma_m(u) = 0, \qquad \ldots(2)$$

where $a_1, a_2, \ldots, a_m$ are all different from 0. If $m = 1$, then the relation (2) reduces to

$$\begin{aligned} & a_1\sigma_1(u) = 0 \ \forall\ u \in K \\ \Rightarrow\ & a_1\sigma_1(1) = 0, \text{ since } 1 \in K \\ \Rightarrow\ & a_1\, 1 = 0, \text{ since } \sigma_1(1) = 1 \\ \Rightarrow\ & a_1 = 0. \end{aligned}$$

But according to our assumption $a_1 \neq 0$. Hence we must have $m > 1$.

Since the automorphisms σ_1 and σ_m are distinct, therefore there exists an element $c \in K$ such that $\sigma_1(c) \neq \sigma_m(c)$.

Now $c \in K$, $u \in K \Rightarrow cu \in K$. Therefore the relation (2) will hold good for cu. Thus we get

$$\begin{aligned} & a_1\sigma_1(cu) + a_2\sigma_2(cu) + \ldots + a_m\sigma_m(cu) = 0 \ \forall\ u \in K \\ \Rightarrow\ & a_1\sigma_1(c)\,\sigma_1(u) + a_2\sigma_2(c)\,\sigma_2(u) + \ldots + a_m\sigma_m(c)\,\sigma_m(u) = 0 \\ & \forall\ u \in K \qquad \ldots(3) \end{aligned}$$

[$\because$ each σ is an automorphism of K]

Multiplying (2) by $\sigma_1(c)$ and subtracting from (3), we get

$$a_2[\sigma_2(c) - \sigma_1(c)]\,\sigma_2(u) + \ldots + a_m[\sigma_m(c) - \sigma_1(c)]\sigma_m(u) = 0. \qquad \ldots(4)$$

Let $b_i = a_i[\sigma_i(c) - \sigma_1(c)]$ for $i = 2, \ldots, m$.

Since a_i, $\sigma_i(c)$, $\sigma_1(c)$ are all in K, therefore each b_i is in K. Further $b_m = a_m[\sigma_m(c) - \sigma_1(c)] \neq 0$, because $a_m \neq 0$, and $\sigma_m(c) - \sigma_1(c) \neq 0$. Therefore from (4), we get

$$b_2\sigma_2(u) + \ldots + b_m\sigma_m(u) = 0 \text{ for all } u \in K. \qquad \ldots(5)$$

The relation (5) has $m-1$ terms. Since $b_m \neq 0$, therefore from (5) we see that it is possible for us to find a relation of the type (1) having less than m non-zero terms. This contradicts the assumption that (2) is a minimal relation.

Hence our initial assumption is wrong. This proves the theorem.

Theorem 35. *Suppose K is a finite extension of a field F. Then show that $G(K, F)$ is a finite group and its order, $o\,(G, (K, F))$ satisfies the relation* $$o(G\,(K, F)) \leqslant [K : F].$$

(G N.D.U. Amritsar 1982; Vikram 76; Calicut 75)

Proof. Suppose $[K : F]=n$ and let $u_1, u_2, \ldots, u_n$ be a basis of the vector space $K(F)$.

Suppose we can find $n+1$ distinct automorphisms $\sigma_1, \sigma_2, \ldots, \sigma_n, \sigma_{n+1}$ in $G(K, F)$. Consider the following system of homogeneous linear equations in $n+1$ unknowns $x_1, x_2, \ldots, x_{n+1}$ and with coefficients in K :

$$\begin{array}{l}
\sigma_1(u_1)\, x_1+\sigma_2(u_1)\, x_2+\ldots+\sigma_{n+1}(u_1)\, x_{n+1}=0\\
\sigma_1(u_2)\, x_1+\sigma_2(u_2)\, x_2+\ldots+\sigma_{n+1}(u_2)\, x_{n+1}=0\\
\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots\\
\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots\\
\sigma_1(u_i)\, x_1+\sigma_2(u_i)\, x_2+\ldots+\sigma_{n+1}\,(u_i)\, x_{n+1}=0\\
\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots\\
\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots\\
\sigma_1(u_n)\, x_1+\sigma_2(u_n)\, x_2+\ldots+\sigma_{n+1}(u_n)\, x_{n+1}=0.
\end{array}$$

Since in the above system of linear homogeneous equations the number of equations is less than the number of unknowns, therefore this system must possess a non-trivial solution (not all 0) $x_1=a_1, x_2=a_2, \ldots, x_{n+1}=a_{n+1}$ in K. Therefore we have,

$$a_1\sigma_1(u_i)+a_2\sigma_2(u_i)+\ldots+a_{n+1}\sigma_{n+1}(u_i)=0, \qquad \ldots(1)$$
$$\text{for} \quad i=1, 2, \ldots, n.$$

Let t be an arbitrary element of K. Since $\{u_1, \ldots, u_n\}$ is a basis of K over F, therefore we can find $\alpha_1, \ldots, \alpha_n \in F$ such that

$$t=\alpha_1 u_1+\ldots+\alpha_n u_n.$$

Now $\quad \sigma_1(t)=\sigma_1(\alpha_1 u_1+\ldots+\alpha_n u_n)$

$=\sigma_1(\alpha_1)\,\sigma_1(u_1)+\ldots+\sigma_1(\alpha_n)\,\sigma_1(u_n)$ [$\because$ σ_1 is an automorphism]

$=\alpha_1\sigma_1(u_1)+\ldots+\alpha_n\sigma_1(u_n)$ [$\because$ $\sigma_1 \in G(K, F) \Rightarrow \sigma_1$ leaves every element of F fixed. Note that $\alpha_1, \ldots, \alpha_n \in F$]

Thus $\quad \sigma_1(t)=\alpha_1\sigma_1(u_1)+\ldots+\alpha_n\sigma_1(u_n)$.

Similarly $\quad \sigma_2(t)=\alpha_1\sigma_2(u_1)+\ldots+\alpha_n\sigma_2(u_n)$

$$\sigma_{n+1}(t)=\alpha_1\sigma_{n+1}(u_1)+\ldots+\alpha_n\sigma_{n+1}(u_n).$$

Multiplying both sides of these equations by $a_1, a_2, \ldots, a_{n+1}$ respectively and adding, we get

Theorem 38. *K is a normal extension of a field F of characteristic 0 if and only if K is the splitting field of some polynomial over F.* (G. N. D. U. Amritsar 1982)

Proof. Suppose K is a normal extension of F. Then K is necessarily a finite extension of F. Also characteristic F is 0. Therefore by theorem 29 K is a simple extension of F. So there is an element a in K such that $K = F(a)$. The group $G(K, F)$ is of finite order say of order n. Let $\sigma_1, \sigma_2, \ldots, \sigma_n$ denote all the distinct elements of $G(K, F)$. Also let σ_1 be the identity of $G(K, F)$ *i.e.*, σ_1 be the identity automorphism of K.

Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be the elementary symmetric functions of the elements $\sigma_1(a), \ldots, \sigma_n(a) \in K$ *i.e.*,

$$\alpha_1 = \sum_{i=1}^{n} \sigma_i(a)$$

$$\sigma_2 = \sum_{i<j} \sigma_i(a)\, \sigma_j(a)$$

$$\cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots$$

$$\cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots$$

$$\alpha_n = \sigma_1(a)\, \sigma_2(a) \ldots \sigma_n(a).$$

It can be easily seen that each α_i is invariant under every $\sigma \in G(K, F)$.

Therefore each α_i is in the fixed field of $G(K, F)$. But K is a normal extension of F implies that F is the fixed field of $G(K, F)$. So each $\alpha_i \in F$. Consider the polynomial

$p(x) = [x - \sigma_1(a)]\,[x - \sigma_2(a)] \ldots [x - \sigma_n(a)]$ over K. ...(1)

We have $p(x) = x^n - \alpha_1 x^{n-1} + \ldots + (-1)^n \alpha_n$.

Since each $\alpha_i \in F$, therefore $p(x) \in F[x]$.

From (1), we see that K splits $p(x) \in F[x]$ into a product of linear factors. Also from (1) we see that $\sigma_1(a) = a$ is a root of $p(x)$ in K. Since a generates K over F, therefore a can be in no proper subfield of K which contains F. Thus K is the splitting field of $p(x)$ over F. This proves the 'only if' part of the theorem.

Now we shall prove the 'if' part of the theorem. Suppose K is the splitting field of the polynomial $f(x) \in F[x]$. Then to prove that K is a normal extension of F. Certainly K is a finite extension of F because every splitting field is a finite extension.

Let $[K : F] = n$. We shall prove the result by induction on n.

To start the induction we see that if $[K : F] = 1$, then $K = F$ and certainly F is a normal extension of F.

Now assume as our induction hypothesis that if K_1 is the splitting field over F_1 of a polynomial in $F_1[x]$ and if $[K_1 : F_1] < n$, then K_1 is a normal extension of F_1.

Now let $[K : F]=n > 1$. K is the splitting field over F of $f(x)$ in $F[x]$. If $f(x)$ splits into linear factors over F, then $K=F$ and consequently $[K : F]=1$. So, assume that $f(x)$ has an irreducible factor $p(x) \in F[x]$ of degree $r > 1$. Since characteristic F is 0 and $p(x)$ is irreducible, therefore $p(x)$ cannot have multiple roots. Further $p(x)$ is a factor of $f(x)$ and K is the splitting field of $f(x)$. Therefore $p(x)$ has a full complement of roots in K. Let $\alpha_1, \alpha_2, \ldots, \alpha_r \in K$ be the r distinct roots of $p(x)$.

Since $p(x)$ is irreducible over F and $\deg p(x)=r$, therefore

$$[F(\alpha_1) : F]=r.$$

We have $[K : F]=[K : F(\alpha_1)]\,[F(\alpha_1) : F]$

$$\Rightarrow [K : F(\alpha_1)]=\frac{[K : F]}{[F(\alpha_1) : F]}=\frac{n}{r} < n.$$

Now K is also the splitting field of $f(x)$ considered as a polynomial over $F(\alpha_1)$.

Since $[K : F(\alpha_1)] < n$, therefore by our induction hypothesis K is a normal extension of $F(\alpha_1)$.

Let $\theta \in K$ be left fixed by every automorphism in $G(K, F)$. If we prove that θ is in F, then F will be the fixed field of $G(K, F)$ and K will be a normal extension of F.

We have

$$\sigma \in G(K, F(\alpha_1))$$

$\Rightarrow$ σ leaves every element of $F(\alpha_1)$ fixed

$\Rightarrow$ σ leaves every element of F fixed [$\because$ $F \subseteq F(\alpha_1)$]

$\Rightarrow$ $\sigma \in G(K, F)$

$\Rightarrow$ $\sigma(\theta)=\theta$ [$\because$ by our assumption θ is left fixed by every automorphism in $G(K, F)$].

Thus we have $\sigma(\theta)=\theta$ for all $\sigma \in G(K, F(\alpha_1))$.

Therefore θ belongs to the fixed field of $G(K, F(\alpha_1))$.

But K is a normal extension of $F(\alpha_1)$ means that $F(\alpha_1)$ is the fixed field of $G(K, F(\alpha_1))$. Therefore θ is in $F(\alpha_1)$. So we can write

$$\theta=\lambda_0+\lambda_1\alpha_1+\lambda_2\alpha_1^2+\ldots+\lambda_{r-1}\,\alpha_1^{r-1} \text{ where } \lambda_0, \ldots, \lambda_{r-1} \in F \quad \ldots(2)$$

[See theorem 5]

Now by theorem 37, for each $i=1, 2, \ldots, r$ there is an automorphism σ_i of K, $\sigma_i \in G(K, F)$, such that $\sigma_i(\alpha_1)=\alpha_i$. Since θ is left fixed by every automorphism in $G(K, F)$, therefore $\sigma_i(\theta)=\theta$. Also $\lambda_0, \lambda_1, \ldots, \lambda_{r-1} \in F$

$$\Rightarrow \sigma_i(\lambda_0)=\lambda_0, \sigma_i(\lambda_1)=\lambda_1, \ldots, \sigma_i(\lambda_{r-1})=\lambda_{r-1}.$$

Operating σ_i on both sides of (2), we get

$$\sigma_i(\theta)=\sigma_i(\lambda_0+\lambda_1\alpha_1+\lambda_2\alpha_1^2+\ldots+\lambda_{r-1}\alpha_1^{r-1})$$

$$\Rightarrow \theta=\sigma_i(\lambda_0)+\sigma_i(\lambda_1)\sigma_i(\alpha_1)+\sigma_i(\lambda_2)[\sigma_i(\alpha_1)]^2+\ldots$$
$$+\sigma_i(\lambda_{r-1})[\sigma_i(\alpha_1)]^{r-1}$$

$$\Rightarrow \theta=\lambda_0+\lambda_1\alpha_i+\lambda_2\alpha_i^2+\ldots+\lambda_{r-1}\alpha_i^{r-1}$$

$$\Rightarrow \lambda_{r-1}\alpha_i^{r-1}+\lambda_{r-2}\alpha_i^{r-2}+\ldots+\lambda_1\alpha_i+(\lambda_0-\theta)=0$$
$$\text{for } i=1, 2, \ldots, r. \qquad \ldots(3)$$

Since $\alpha_1, \alpha_2 \ldots, \alpha_r$ are all distinct, therefore from (3) we see that the polynomial

$$q(x)=\lambda_{r-1}x^{r-1}+\lambda_{r-2}x^{r-2}+\ldots+\lambda_1 x+(\lambda_0-\theta)$$

in $K[x]$, of degree at most $r-1$, has r distinct roots $\alpha_1, \alpha_2, \ldots, \alpha_r$. This can be possible only if all the coefficients of $q(x)$ are 0; in particular $\lambda_0-\theta=0$ which gives $\theta=\lambda_0$. So θ is in F.

Hence K is a normal extension of F.

The proof of the theorem is now complete by induction.

Note. The above theorem is very important. It tells us that if K is the splitting field of some polynomial $f(x) \in F[x]$ where characteristic F is 0, then K is a normal extension of F. The converse is also true.

Theorem 39. *Let K be a normal extension of a field F of characteristic* 0. *If T is a subfield of K containing F, then T is a normal extension of F if and only if*

$$\sigma(T) \subseteq T \ \forall \ \sigma \in G(K, F).$$

Proof. Since K is a normal extension therefore K is a finite extension of F. Let T be a subfield of K containing F. Then T is also a finite extension of F which is of characteristic 0. Therefore by theorem 29, T is a simple extension of F. So there is an element a in T such that $T=F(a)$.

'Only if' part of the theorem. Suppose $\sigma(T) \subseteq T \forall \sigma \in G(K, F)$. Then to prove that T is a normal extension of F.

Since K is a normal extension of F, therefore $G(K, F)$ is a finite group say of order n. Let $\sigma_1, \sigma_2, \ldots, \sigma_n$ be all the n elements of $G(K, F)$ and let σ_1 be the identity of $G(K, F)$. Since

$$a \in T \text{ and } \sigma(T) \subseteq T \ \forall \ \sigma \in G(K, F),$$

therefore $\sigma_1(a), \sigma_2(a), \ldots, \sigma_n(a)$ are all in T. Consider the polynomial $\quad p(x)=[x-\sigma_1(a)][x-\sigma_2(a)]\ldots[x-\sigma_n(a)]$

over T. Expanding we get

$p(x)=x^n-\alpha_1x^{n-1}+\alpha_2x^{n-2}+\ldots+(-1)^n\alpha_n$ where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are elementary symmetric functions of the elements

$$\sigma_1(a), \sigma_2(a), \ldots, \sigma_n(a)$$

of T. It can be easily seen that $\alpha_1, \alpha_2, \ldots, \alpha_n$ are each invariant with respect to every $\sigma \in G\,(K, F)$. Therefore each α_i for $i=1, \ldots, n$ must be an element of the fixed field of $G\,(K, F)$. Since K is a normal extension of F, therefore the fixed field of $G\,(K, F)$ is F. Thus each α_i is in F and so $p(x) \in F[x]$. From the form of $p(x)$ we see that T splits the polynomial $p(x) \in F[x]$ into a product of linear factors. Also $\sigma_1(a)=a$ is a root of $p(x)$ in $T=F(a)$. Since a generates T over F, therefore a can be in no proper subfield of T which contains F. Thus T is the splitting field of $p(x) \in F[x]$. Therefore by theorem 38, T is a normal extension of F.

'If' part of the theorem. Suppose T is a normal extension of F. Then to prove that $\sigma(T) \subseteq T\ \forall\ \sigma \in G\,(K, F)$.

Since T is a normal extension of F, therefore $G\,(T, F)$ is a finite group say of order m. Let $\psi_1, \psi_2, \ldots, \psi_m$ be all the m elements of $G(T, F)$ and let ψ_1 be the identity of $G\,(T, F)$. Since $a \in T=F(a)$, therefore $\psi_1(a), \psi_2(a), \ldots, \psi_m(a)$ are all in T. Consider the polynomial $q(x)=[x-\psi_1(a)]\,[x-\psi_2(a)]\ldots[x-\psi_m(a)]$ over T. Expanding we get $q(x)=x^m-\beta_1x^{m-1}+\beta_2x^{m-2}-\ldots+(-1)^m\beta_m$ where $\beta_1, \beta_2, \ldots, \beta_m$ are elementary symmetric functions of the elements $\psi_1(a), \psi_2(a), \ldots, \psi_m(a)$ of T. It can be easily seen that each β_i is invariant with respect to each ψ in $G(T, F)$. Since T is a normal extension of F, therefore each β_i must be in F. Consequently $q(x) \in F[x]$. From the form of $q(x)$ we see that $q(x)$ has all its roots in T. Also $\psi_1(a)=a$ is a root of $q(x)$ in T and hence in K.

Now let σ be any element of $G(K, F)$. Since σ is an automorphism of K leaving every element of F fixed and a is a root of $q(x) \in F[x]$ in K, therefore $\sigma(a)$ is also a root of $q(x)$ in K. Since all roots of $q(x)$ are in T, therefore $\sigma(a)$ must be an element of T. Now $T=F(a)$. If $[T : F]=r$, then $t \in T$ can be written as

$t=\lambda_0+\lambda_1a+\ldots+\lambda_{r-1}a^{r-1}$ where $\lambda_0, \lambda_1, \ldots, \lambda_{r-1} \in F$.

We have $\sigma(t)=\sigma(\lambda_0+\lambda_1a+\ldots+\lambda_{r-1}a^{r-1})$

$=\sigma(\lambda_0)+\sigma(\lambda_1)\,\sigma(a)+\ldots+\sigma\,(\lambda_{r-1})\,[\sigma(a)]^{r-1}$

$=\lambda_0+\lambda_1\sigma(a)+\ldots+\lambda_{r-1}\,[\sigma(a)]^{r-1}$

[$\because$ σ leaves every element of F fixed]

$$\begin{aligned}&a_1\sigma_1(t)+\ldots+a_{n+1}\sigma_{n+1}(t)\\ &=\alpha_1[a_1\sigma_1(u_1)+\ldots+a_{n+1}\sigma_{n+1}(u_1)]+\ldots+\alpha_n[a_1\sigma_1(u_n)+\ldots\\ &\qquad\qquad+a_{n+1}\sigma_{n+1}(u_n)]\\ &=\alpha_1 0+\ldots+\alpha_n 0, \text{ by virtue of the system of equations (1)}\\ &=0.\end{aligned}$$

Thus we see that if $\sigma_1, \ldots, \sigma_{n+1}$ are distinct automorphisms of K, then it is possible to find $a_1, \ldots, a_{n+1}$ in K, not all 0, such that

$$a_1\sigma_1(t)+\ldots+a_{n+1}\sigma_{n+1}(t)=0 \ \forall \ t \in K.$$

By theorem 34 this is not possible for us. Hence there cannot be $n+1$ distinct automorphisms in $G(K, F)$.

Similarly we can prove that there cannot be $n+2, n+3, \ldots$ distinct automorphisms in $G(K, F)$. Hence $o(G(K, F)) \leqslant n$.

Elementary symmetric functions of the elements of a field.

Definition. *Let* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *be* n *elements of a field* F. *Then*

$$y_1=\alpha_1+\alpha_2+\ldots+\alpha_n=\sum_{i=1}^{n} \alpha_i$$

$$y_2=\sum_{i<j} \alpha_i \alpha_j$$

$$y_3=\sum_{i<j<k} \alpha_i\alpha_j\alpha_k$$

$$y_n=\alpha_1\,\alpha_2\,\alpha_3\ldots\alpha_n$$

are called elementary symmetric functions of $\alpha_1, \alpha_2, \ldots, \alpha_n$.

It can be easily seen by direct calculation that

$$\begin{aligned}&(x-\alpha_1)\,(x-\alpha_2)\ldots(x-\alpha_n)\\ &\quad=x^n-y_1x^{n-1}+y_2x^{n-2}-y_3x^{n-3}+\ldots+(-1)^n\,y_n.\end{aligned}$$

Normal extension of a field. Definition. *A finite extension K of a field F is said to be a normal extension of F if the fixed field of $G(K, F)$ is F.* **(Meerut 1981, 82, 83, 84P)**

From the definition of normal extension it is obvious that if a is any element of K which is not in F, then we must have some automorphism σ in $G(K, F)$ such that $\sigma(a) \neq a$.

Theorem 36. *Suppose K is a finite extension of a field F of characteristic* 0 *and H is a subgroup of $G(K, F)$. Let K_H be the fixed field of H. Then*

(1) $[K : K_H]=o(H)$. (2) $H=G(K, K_H)$.

Proof. Let $A(K)$ be the group of all automorphisms of the field K which is a finite extension of the field F. It is given that K_H is the fixed field of H which is a subgroup of $G(K, F)$. Therefore

$$K_H=\{x\in K : \sigma(x)=x \text{ for all } \sigma\in H\}.$$

Now K is a finite extension of F and K_H is a subfield of K containing F. Therefore K is a finite extension of K_H *i.e.*, $[K : K_H]$ is finite. By theorem 35, we have $G(K, K_H)$ is a finite group and

$$o(G(K, K_H)) \leqslant [K : K_H]. \qquad \text{...(1)}$$

Further $G(K, K_H)$ is a subgroup of $A(K)$ and H is also a subgroup of $A(K)$. We have

$\sigma\in H \Rightarrow \sigma(y)=y \; \forall \; y \in K_H$ [see def. of K_H]

$\Rightarrow \sigma \in G(K, K_H)$.

$\therefore \quad H \subseteq G(K, K_H)$

$\Rightarrow H$ is a subgroup of $G(K, K_H)$

$\Rightarrow o(H) \leqslant o(G(K, K_H))$. ...(2)

From (1) and (2), we get

$$o(H) \leqslant o[G(K, K_H)] \leqslant [K : H_H] \qquad \text{...(3)}$$

$$\Rightarrow o(H) \leqslant [K : K_H] \qquad \text{...(4)}$$

Now in order to prove that $o(H)=[K : K_H]$ it remains to prove that $o(H) \geqslant [K : K_H]$.

Let $o(H)=h$ and $[K : K_H]=m$.

Let $H=\{\sigma_1, \sigma_2, \ldots, \sigma_h\}$ where σ_1 is the identity automorphism of K. Note that H being a subgroup of $A(K)$ must contain its identity.

Since K is a finite extension of K_H which is of characteristic 0, therefore by theorem 29, K is a simple extension of K_H. Therefore there exists an element $a\in K$ such that $K=K_H(a)$. This a must therefore satisfy an irreducible polynomial over K_H of degree $m=[K : K_H]$ and no non-trivial polynomial of lower degree.

[see theorem 7].

Let $\alpha_1, \alpha_2, \ldots, \alpha_h$ be the elementary symmetric functions of the elements $\sigma_1(a), \sigma_2(a), \ldots, \sigma_h(a) \in K$. Then

$$\alpha_1=\sigma_1(a)+\sigma_2(a)+\ldots+\sigma_h(a)=\sum_{i=1}^{h} \sigma_i(a)$$

$$\alpha_2=\sum_{i<j} \sigma_i(a)\, \sigma_j(a)$$

$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$

$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$

$$\alpha_h=\sigma_1(a)\, \sigma_2(a) \cdots \sigma_h(a).$$

We claim that each α_i is invariant under every $\sigma \in H$. Let σ be any element of H. Consider the h products $\sigma\sigma_1, \sigma\sigma_2, \ldots, \sigma\sigma_h$.

All these are elements of H because H is a group. Also all these are distinct. Because $\sigma\sigma_i = \sigma\sigma_j$,

$$\Rightarrow \sigma_i = \sigma_j \quad \text{[by left cancellation law in } H]$$

$\therefore$ $\sigma\sigma_1, \sigma\sigma_2, \ldots, \sigma\sigma_h$ are nothing but the h elements of H placed in some order. We have

$$\sigma(\alpha_1) = \sigma\,[\sigma_1(a) + \sigma_2(a) + \ldots + \sigma_h(a)]$$

$= \sigma\,[\sigma_1(a)] + \sigma\,[\sigma_2(a)] + \ldots + \sigma\,[\sigma_h(a)]$ $\quad$ [$\because$ σ is an automorphism]

$= (\sigma\sigma_1)(a) + (\sigma\sigma_2)(a) + \ldots + (\sigma\sigma_h)(a)$

$= \sum_{i=1}^{h} \sigma_i(a)$, since $\sigma\sigma_1, \ldots, \sigma\sigma_h$ are nothing but the h elements of H placed in some order

$= \alpha_1$.

Further $\sigma(\alpha_2) = \sigma \sum_{i<j} \sigma_i(a)\,\sigma_j(a)$

$= \sum_{i<j} \sigma\,[\sigma_i(a)\,\sigma_j(a)]$ $\quad$ [$\because$ σ is an automorphism]

$= \sum_{i<j} \sigma\,[\sigma_i(a)]\,\sigma\,[\sigma_j(a)]$ $\quad$ [$\because$ σ is an automorphism]

$= \sum_{i<j} (\sigma\sigma_i)(a)\,(\sigma\sigma_j)(a) = \alpha_2$

Similarly we can show for $\alpha_3, \alpha_4, \ldots, \alpha_h$. Since each α_i is invariant under every $\sigma \in H$, therefore by the definition of K_H each $\alpha_i \in K_H$. Consider the polynomial

$$p(x) = [x - \sigma_1(a)]\,[x - \sigma_2(a)] \ldots [x - \sigma_h(a)] \qquad \ldots(5)$$

$$= x^h - \alpha_1 x^{h-1} + \alpha_2 x^{h-2} - \ldots + (-1)^h \alpha_h.$$

We see that $p(x)$ is a non-zero polynomial over K_H of degree h. Note that $\alpha_1, \ldots, \alpha_h \in K_H$. Also from (5) we see that $\sigma_1(a) = a$ is a root of $p(x)$. Note that $\sigma_1(a) = a$ because σ_1 is identity automorphism.

Thus we see that a satisfies a non-trivial polynomial over K_H of degree h. Since a cannot satisfy a non-trivial polynomial over K_H of degree lower than m, therefore we must have

$$h \geq m$$

$$\Rightarrow o(H) \geq [K : K_H]. \qquad \ldots(6)$$

From (4) and (6), we get $o(H) = [K : K_H]$.

Putting $o(H) = [K : K_H]$ in (3), we get

$$o(H) \leq o(G(K, K_H)) \leq o(H)$$

$$\Rightarrow o(H) = o(G(K, K_H)).$$

Now H is a subgroup of $G(K, K_H)$.

Therefore $o(H)=o(G(K, K_H))$ implies that $H=G(K, K_H)$.

Corollary. *Let K be a normal extension of a field F of characteristic 0. Then* $[K : F]=o(G(K, F))$.

Proof. Since K is normal extension of F, therefore the fixed field of $G(K, F)$ is F itself. Also K is a finite extension of F because a normal extension is necessarily a finite extension.

Now in the proof of the above theorem take $H=G(K, F)$ itself. Then $K_H=$the fixed field of $G(K, F)=F$. Therefore we get the result $o(G(K, F))=[K : F]$.

Note. Write the complete proof of this corollary by replacing H by $G(K, F)$ and K_H by F in the proof of the above theorem.

Now we shall prove a very important theorem which characterizes normal extensions of a field of characteristic 0. Before stating the main theorem we want to prove the following subsidiary theorem which we shall use in the proof of the main theorem.

Theorem 37. *Let K be the splitting field of* $f(x)$ *in* $F[x]$ *and let* $p(x)$ *be an irreducible factor of* $f(x)$ *in* $F[x]$. *If the roots of* $p(x)$ *are* $\alpha_1, \ldots, \alpha_r$, *then for each i there exists an automorphism* σ_i *in* $G(K, F)$ *such that* $\sigma_i(\alpha_1)=\alpha_i$.

Proof. It is given that K is the splitting field of $f(x) \in F[x]$ and $p(x)$ is an irreducible factor of $f(x)$ in $F[x]$. Since $p(x)$ is a factor of $f(x)$, therefore every root of $p(x)$ is also a root of $f(x)$ and so every root of $p(x)$ must be in K. Let α_1, α_i be any two roots of $p(x)$. Let $F_1=F(\alpha_1)$ and $F_1'=F(\alpha_i)$. Since $p(x) \in F[x]$ is irreducible, therefore by corollary to theorem 21, there is an isomorphism ψ of F_1 onto F_1' such that

$$\psi(\alpha_1)=\alpha_i \text{ and } \psi(\alpha)=\alpha \ \forall\ \alpha \in F.$$

Since F_1 is a superfield of F, therefore $f(x) \in F[x]$ can be considered as a polynomial over F_1. We claim that K is a splitting field for $f(x) \in F_1[x]$ because no subfield of K, containing F_1 and hence F, can split $f(x)$. Similarly K is a splitting field for $f(x)$ considered as a polynomial over F_1'. By theorem 22, there is an isomorphism σ_i of K onto K (thus an automorphism of K) coinciding with ψ on F_1.

Now $\alpha_1 \in F(\alpha_1)=F_1$

$$\Rightarrow \sigma_i(\alpha_1)=\psi(\alpha_1)=\alpha_i. \qquad [\because \ \psi(\alpha_1)=\alpha_i]$$

Also since σ_i coincides with ψ on F_1, therefore $\sigma_i(\alpha)=\alpha \ \forall\ \alpha \in F$. Consequently $\sigma_i \in G(K, F)$.

This proves the theorem.

i.e., if and only if $(\sigma^{-1}\tau\sigma)(t)=t \;\forall\; t\in T,\;\forall\;\sigma\in G(K,F)$
and $\forall\;\tau\in G(K,T)$

i.e., if and only if $\sigma^{-1}\tau\sigma\in G(K,T)\;\forall\;\sigma\in G(K,F)$
and $\forall\;\tau\in G(K,T)$

[Note that $\sigma^{-1}\tau\sigma\in G(K,T)\Leftrightarrow(\sigma^{-1}\tau\sigma)(t)=t\;\forall\;t\in T$]

i.e., if and only if $G(K,T)$ is a normal subgroup of $G(K,F)$
[by def of normal subgroup].

(3) If T is a normal extension of F, then for any $\sigma\in G(K,F)$, we have

$$\sigma(T)\subseteq T$$
$$\Rightarrow \sigma(t)\in T\;\forall\;t\in T.$$

Therefore σ induces an automorphism σ^* of T defined as

$$\sigma^*(t)=\sigma(t)\;\forall\;t\in T.$$

Since σ leaves every element of F fixed, therefore σ^* also leaves every element of F fixed. So σ^* must be in $G(T,F)$.

If $\sigma_1,\sigma_2\in G(K,F)$, then $\forall\;t\in T$, we have

$$\begin{aligned}(\sigma_1\sigma_2)^*(t)&=(\sigma_1\sigma_2)(t)\\&=\sigma_1[\sigma_2(t)]=\sigma_1[\sigma_2^*(t)]=\sigma_1^*[\sigma_2^*(t)]\\&=(\sigma_1^*\sigma_2^*)(t).\end{aligned}$$

$\therefore\quad(\sigma_1\sigma_2)^*=\sigma_1^*\sigma_2^*.$

From this we conclude that the mapping ϕ of $G(K,F)$ into $G(T,F)$ defined as $\phi(\sigma)=\sigma^*$ for all $\sigma\in G(K,F)$ is a homomorphism of $G(K,F)$ into $G(T,F)$. The Kernel of this homomorphism consists of elements σ in $G(K,F)$ such that $\phi(\sigma)=\sigma^*$ is the identity of the group $G(T,F)$. The identity of the group $G(T,F)$ is the identity map on T. Therefore the Kernel of ϕ is the set of all $\sigma\in G(K,F)$ such that $t=\sigma^*(t)=\sigma(t)$ for all $t\in T$. But $\sigma(t)=t\;\forall\;t\in T$ if and only if $\sigma\in G(K,T)$. Therefore the kernel of ϕ is exactly $G(K,T)$. Now by the fundamental theorem on homomorphism of groups the image of $G(K,F)$ in $G(T,F)$ under the mapping ϕ is isomorphic to the quotient group

$$G(K,F)/G(K,T).$$

Now the order of the group $G(K,F)/G(K,T)$

$$\begin{aligned}&=\frac{o(G(K,F))}{o(G(K,T))}\\&=\text{the index of } G(K,T) \text{ in } G(K,F)\\&=[T:F], \text{ by part (1) of this theorem}\\&=o(G(T,F))\qquad\text{[by corollary to theorem 26}\\&\qquad\qquad\qquad\text{since } T \text{ is normal over } F]\end{aligned}$$

Thus the image of $G(K, F)$ in $G(T, F)$ is isomorphic to a group of order $o(G(T, F))$. Since the image of $G(K, F)$ in $G(T, F)$ is a subgroup of $G(T, F)$, therefore it is all of $G(T, F)$. Hence the quotient group $G(K, F)/G(K, T)$ is isomorphic to $G(T, F)$.

Solved Examples

Ex. 1. *If K is a field and S a set of automorphisms of K, prove that the fixed field of S and that of $\bar{S}$, (the subgroup of the group of all automorphisms of K generated by S) are identical.*

Solution. Let S be a set of automorphisms of a field K and $\bar{S}$ be the subgroup of the group of all automorphisms of K generated by S. Let L_1 and L_2 be the fixed fields of S and S respectively. To prove that $L_1 = L_2$.

We have
$$\begin{aligned} a \in L_2 &\Rightarrow \sigma(a) = a \ \forall\ \sigma \in \bar{S} \\ &\Rightarrow \sigma(a) = a \ \forall\ \sigma \in S \\ &\qquad [\because\ S \subseteq \bar{S} \Rightarrow \text{every } \sigma \in S \text{ is also in } \bar{S}] \\ &\Rightarrow a \in L_1. \end{aligned}$$

$\therefore\ L_2 \subseteq L_1$.

If σ is an automorphism of K, then $\sigma(a) = a \Leftrightarrow \sigma^{-1}(a) = a$. Consequently if $\sigma(a) = a$, then $\sigma^n(a) = a$ where n is any integer. Note that $\sigma^2(a) = \sigma\,[\sigma(a)] = \sigma(a) = a$ and so on.

Now let $a \in L_1$ be arbitrary. Then $\sigma(a) = a \ \forall\ \sigma \in S$.

Let f be any element of $\bar{S}$. Since $\bar{S}$ is a subgroup of the group of all automorphisms of K generated by S, therefore f can be expressed as a product of positive and negative integral powers of a finite number of elements of S Let $f = \sigma_1^{p_1}\, \sigma_2^{p_2} \ldots \sigma_m^{p_m}$ where $\sigma_1, \ldots, \sigma_m \in S$ and $p_1, \ldots, p_m$ are any integers.

We have $f(a) = (\sigma_1^{p_1} \ldots \sigma_m^{p_m})\,(a)$

$= a\ [\because\ \sigma_1, \ldots, \sigma_m \in S \Rightarrow \sigma_i(a) = a \text{ for all } i = 1, \ldots, m]$.

Thus $a \in L_1 \Rightarrow f(a) = a\ \forall\ f \in \bar{S} \Rightarrow a \in L_2$.

$\therefore\ L_1 \subseteq L_2$ Hence $L_1 = L_2$.

Ex. 2. *Let K be an extension of the field of rational numbers F. Show that any automorphism of K must leave every element of F fixed.*

Solution. Let σ be any automorphism of K. If p is any integer we claim that $\sigma(p) = p$.

Case 1. p is 0. In this case $\sigma(p) = \sigma(0) = 0 = p$.

Case 2. p is a positive integer.

We have $\quad p = 1 + 1 + 1 + \ldots$ upto p times

$$\begin{aligned} \Rightarrow \sigma(p) &= \sigma(1+1+1+\ldots\text{upto } p \text{ times}) \\ &= \sigma(1)+\sigma(1)+\ldots\text{upto } p \text{ times} \\ &= 1+1+\ldots\text{upto } p \text{ times} \qquad [\because \sigma(1)=1] \\ &= p. \end{aligned}$$

Case 3. p is a negative integer, say $p=-q$ where q is a positive integer.

We have $\sigma(p)=\sigma(-q)=\sigma\,[(-1)+(-1)+\ldots\text{upto } q \text{ times}]$

$$=\sigma(-1)+\sigma(-1)+\ldots\text{upto } q \text{ times}=(-1)+(-1)+\ldots\text{upto } q \text{ times} \qquad [\because \sigma(-1)=-\sigma(1)=-1]$$

$$=-q=p.$$

Now let $\alpha \in F$ *i.e.*, let α be any rational number. Then $\alpha=m/n$ where m and n are integers and $n \neq 0$. If σ is any automorphism of K, we have

$$\begin{aligned} \sigma(\alpha) &= \sigma(m/n)=\sigma(mn^{-1})=\sigma(m)\,\sigma(n^{-1})=m\,[\sigma(n)]^{-1} \\ &= mn^{-1}=m/n=\alpha. \end{aligned}$$

Hence σ leaves every element of F fixed.

Ex. 3. *Let F be the field of rational numbers and let $K=F(2^{1/3})$ where $2^{1/3}$ is the real cube root of 2. Show that the only automorphism of K is the identity automorphism. Is K a normal extension of F?*

Solution. K is an extension of the field of rational numbers F. Therefore any automorphism of K will leave every element of F fixed. We have $K=F(2^{1/3})=\{\alpha_0+\alpha_1\, 2^{1/3}+\alpha_2\,(2^{1/3})^2 : \alpha_0, \alpha_1, \alpha_2 \in F\}$.

Let σ be any automorphism of K.

Since $2^{1/3} \in K$, therefore $\sigma(2^{1/3}) \in K$. We have

$$\begin{aligned} [\sigma(2^{1/3})]^3 &= [\sigma(2^{1/3})]\,[\sigma(2^{1/3})]\,[\sigma(2^{1/3})] \\ &= \sigma(2^{1/3}.2^{1/3}.2^{1/3}) \qquad [\because \sigma \text{ is an automorphism}] \\ &= \sigma[(2^{1/3})^3]=\sigma(2)=2, \text{ since } 2 \in F. \end{aligned}$$

Thus $[\sigma(2^{1/3})]^3=2$

$\Rightarrow \sigma(2^{1/3})$ must also be a cube root of 2 lying in K.

But there is only one real cube root of 2 and K is a subfield of the real field. Therefore we must have $\sigma(2^{1/3})=2^{1/3}$.

Now let $\alpha_0+\alpha_1\, 2^{1/3}+\alpha_2\,(2^{1/3})^2$ be any element of K where α_0, α_1, α_2 are rational numbers. We have

$$\begin{aligned} \sigma\,[\alpha_0+\alpha_1\, 2^{1/3}+\alpha_2\,(2^{1/3})^2] &= \sigma\,(\alpha_0+\alpha_1\, 2^{1/3}+\alpha_2\, 2^{1/3}\, 2^{1/3}) \\ &= \sigma(\alpha_0)+\sigma(\alpha_1)\,\sigma(2^{1/3})+\sigma(\alpha_2)\,\sigma(2^{1/3})\,\sigma(2^{1/3}) \\ &= \alpha_0+\alpha_1\, 2^{1/3}+\alpha_2\, 2^{1/3}\, 2^{1/3} \qquad [\because \alpha_0, \alpha_1, \alpha_2 \text{ are rational numbers}] \\ &= \alpha_0+\alpha_1\, 2^{1/3}+\alpha_2\,(2^{1/3})^2. \end{aligned}$$

Hence σ is the identity automorphism of K.

Since identity automorphism is the only automorphism of K, therefore the group $G(K, F)$ consists of only one element *i.e.*, the identity map I of K. If $a \in K$ be arbitrary, we have $I(a)=a$ Therefore the fixed field of $G(K, F)$ is not F but is K. Since the fixed field of $G(K, F)$ is not F, therefore K is not a normal extension of F.

Ex. 4. *Let K be the field of complex numbers and F be the field of real numbers. Show that K is a normal extension of F.*

Solution. We have $K=\{a+ib : a, b \text{ are real numbers}\}$.

K is an extension of F. The elements $1, i$ of K form a basis of K over F. Therefore $[K : F]=2$ and K is a finite extension of F. Let us compute $G(K, F)$. If σ is any automorphism of K, then

$$[\sigma(i)]^2=\sigma(i)\,\sigma(i)=\sigma(i.i)=\sigma(i^2)=\sigma(-1)=-\sigma(1)=-1.$$

$\therefore$ If σ is any automorphism of K, then

$$[\sigma(i)]^2=-1 \Rightarrow \sigma(i)=\pm\sqrt{-1}=\pm i.$$

Now let σ be any automorphism of K and let $\sigma \in G(K, F)$ *i.e.*, σ leaves every real number fixed. If $a+ib$ is any element of K where a, b are real numbers, then

$$\sigma(a+ib)=\sigma(a)+\sigma(i)\,\sigma(b)$$

$$=a\pm ib \qquad [\because \sigma(i)=\pm i \text{ and } a, b \in F \Rightarrow \sigma(a)=a, \sigma(b)=b].$$

Thus if $\sigma \in G(K, F)$, then we must have $\sigma(a+ib)=a\pm ib$.

Now it can be easily seen that each of these possibilities, namely each of the mappings

$$\sigma_1(a+ib)=a+ib \text{ and } \sigma_2(a+ib)=a-ib$$

defines an automorphism of K, σ_1 being the identity automorphism and σ_2 complex-conjugation.

Therefore the group $G(K, F)$ consists of two elements σ_1, σ_2. Thus $o(G(K, F))=2$ Now let us find the fixed field of $G(K, F)$. The fixed field of $G(K, F)$ necessarily contains F. Let $a+ib \in$ the fixed field of $G(K, F)$. Then we must have

$$\sigma(a+ib)=a+ib \;\forall\; \sigma \in G(K, F)$$

$$\Rightarrow \sigma_2(a+ib)=a+ib \qquad [\because \sigma_2 \in G(K, F)]$$

$$\Rightarrow a-ib=a+ib \Rightarrow 2ib=0 \Rightarrow b=0 \Rightarrow a+ib=a \Rightarrow a+ib \in F.$$

Thus the elements of F are the only elements of K which belong to the fixed field of $G(K, F)$. Hence the fixed field of $G(K, F)$ is exactly F. Consequently K is a normal extension of F.

Ex. 5. *Let x^3-2 be a polynomial over the field F_0 of rational numbers.*

$=$an element of T since $\sigma(a)\in T$

and $\lambda_0, \lambda_1, \ldots, \lambda_{r-1}\in T\supseteq F$.

Thus for every $\sigma\in G(K, F)$ and for every $t\in T$, we have $\sigma(t)\in T$. Consequently $\sigma(T)\subseteq T\ \forall\ \sigma\in G(K, F)$.

Galois group of a polynomial over a field.

Definition. *Let $f(x)$ be a polynomial in $F[x]$ and let K be its splitting field over F. The Galois group of $f(x)$ is the group $G(K, F)$ of all those automorphisms of K which leave every element of F fixed.*

Note. If K is an extension of a field F, then an automorphism of K which leaves every element of F fixed is also called an F-automorphism of K. If K is a normal extension of a field F of characteristic 0, then K is the splitting field of some polynomial $f(x)$ in $F[x]$. The group $G(K, F)$ of F-automorphisms of K is also called the Galois group of K over F.

Theorem 40. *The Galois group of a polynomial over a field is isomorphic to a group of permutations of its roots.*

Proof. Let $f(x)$ be a polynomial over the field F. Let K be the splitting field of $f(x)$ over F. Then K is a normal extension of F. Therefore the Galois group $G(K, F)$ of $f(x)$ is of finite order $[K:F]=n$, say. Let $\sigma_1, \sigma_2, \ldots, \sigma_n$ be the n distinct elements of $G(K, F)$. Suppose $f(x)$ has only m distinct roots in K. It is possible that $f(x)$ may have multiple roots. Let $S=\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ be the set of m distinct roots of $f(x)$ in K. Let P be the set of all those permutations on S which change only those elements of S which are not in F *i.e.*, which leave every element of F (if it is any) fixed. If g_1, g_2 are two elements of P, then g_1g_2 will also be an element of P because g_1g_2 will also leave every element of F fixed. Thus P is closed with respect to product of permutations. Therefore P is a subgroup of the group of all permutations on S. We shall prove that the group $G(K, F)$ is isomorphic to the group P.

Let $\sigma\in G(K, F)$. Let σ^* be the restriction of σ to S *i.e.*, $\sigma^*(\alpha)=\sigma(\alpha)\ \forall\ \alpha\in S$. If α is a root of $f(x)$ in K, then $\sigma(\alpha)$ is also a root of $f(x)$ in K [See Ex. 5 Page 514]. Therefore $\alpha\in S\Rightarrow\sigma(\alpha)=\sigma^*(\alpha)\in S$. So σ^* is a function from S to S. Further σ^* is one-one since σ is one-one. Since S is a finite set, therefore σ^* is one-one implies that σ^* is also onto. Thus σ^* is a permutation on S. Since σ leaves every element of F fixed, therefore σ^* will also leave every element of F fixed. Thus $\sigma^*\in P$.

Now let ψ be a mapping from $G(K, F)$ to P defined as

$$\psi(\sigma)=\sigma^* \ \forall \ \sigma \in G(K, F)$$

ψ is one-one. Let $\sigma_1, \sigma_2 \in G(K, F)$ be arbitrary. Then

$$\begin{aligned} &\psi(\sigma_1)=\psi(\sigma_2) \\ &\Rightarrow \sigma_1^*=\sigma_2^* \\ &\Rightarrow \sigma_1^*(\alpha)=\sigma_2^*(\alpha) \ \forall \ \alpha \in S \\ &\Rightarrow \sigma_1(\alpha)=\sigma_2(\alpha) \ \forall \ \alpha \in S. \end{aligned}$$

Now $K=F(\alpha_1, \alpha_2, \ldots, \alpha_m)$. Therefore each element of K is obtainable as the result of a finite number of operations of addition, subtraction, multiplication and division performed on the elements of F and on $\alpha_1, \ldots, \alpha_m$. But σ_1, σ_2 are automorphisms of K. Each of them leaves every element of F fixed and each of them maps every element of S identically.

Hence
$$\begin{aligned} &\sigma_1(\alpha)=\sigma_2(\alpha) \ \forall \ \alpha \in S \\ &\Rightarrow \sigma_1(\beta)=\sigma_2(\beta) \ \forall \ \beta \in K \\ &\Rightarrow \sigma_1=\sigma_2 \Rightarrow \psi \text{ is one-one.} \end{aligned}$$

ψ is onto.

Let g be any element of P. Then g is a peremutation on S leaving those elements of S fixed which are in F. Now suppose there exists an automorphism σ of K which leaves every element of F fixed and which maps each element of S in the same way as g maps it. Since each element of K is obtainable as the result of a finite number of operations of addition, subtraction, multiplication and division performed on the elements of F and on $\alpha_1, \ldots, \alpha_m$, therefore the automorphism σ of K will be completely determined.

Also the restriction of σ to S will be g. Thus

$$g \in P \Rightarrow \exists \ \sigma \in G(K, F) \text{ such that } \psi(\sigma)=g.$$

Hence ψ is onto.

ψ preserves compositions. Let $\sigma_1, \sigma_2 \in G(K, F)$ be arbitrary. Then

$\psi(\sigma_1\sigma_2)=(\sigma_1\sigma_2)^*$, the restriction of $\sigma_1\sigma_2$ to S.

But $\forall \ \alpha \in S$, we have

$$\begin{aligned} (\sigma_1\sigma_2)^*(\alpha)&=(\sigma_1\sigma_2)(\alpha) && [\text{by def. of } \sigma^*] \\ &=\sigma_1[\sigma_2(\alpha)] \\ &=\sigma_1[\sigma_2^*(\alpha)] && [\text{by def. of } \sigma^*] \\ &=\sigma_1^*[\sigma_2^*(\alpha)]=(\sigma_1^*\sigma_2^*)(\alpha). \end{aligned}$$

$\therefore \ (\sigma_1\sigma_2)^*=\sigma_1^*\sigma_2^*.$

$\therefore \ \psi(\sigma_1\sigma_2)=\sigma_1^*\sigma_2^*=\psi(\sigma_1)\psi(\sigma_2).$

Hence ψ is an isomorphic mapping from $G(K, F)$ onto P.

Therefore $G(K, F)$ is isomorphic to P.

Now we come to the fundamental theorem of Galois theory. It sets up a one-to-one correspondence between the subfields of the splitting field K of $f(x) \in F[x]$ and the subgroups of its Galois group $G(K, F)$.

Fundamental theorem of Galois theory.

Theorem 41. *Let K be a normal extension of a field F, of characteristic 0. Show that there is a one-to-one correspondence between the set of subfields of K which contain F and the set of subgroups of $G(K, F)$. Further, if T is any subfield of K which contains F, then*

(1) *$[K:T]=o(G(K, T))$, and $[T:F]=$ index of $G(K, T)$ in $G(K, F)$.*

(2) *T is a normal extension of F if and only if $G(K, T)$ is a normal subgroup of $G(K, F)$.*

(3) *If T is a normal extension of F, then $G(T, F)$ is isomorphic to $G(K, F)/G(K, T)$.*

(Delhi 1970; Nagarjuna 78, Meerut 80, 81, 82, 83, 84 P; Indore 70)

Proof. For any subfield T of K which contains F, let $G(K, T)$ be the group of all those automorphisms of K which leave every element of T fixed. We have

$\sigma \in G(K, T) \Rightarrow \sigma$ leaves every element of T fixed

$\Rightarrow \sigma$ leaves every element of F fixed, since $F \subseteq T$

$\Rightarrow \sigma \in G(K, F)$.

Thus $G(K, T) \subseteq G(K, F)$. Since both $G(K, F)$ and $G(K, T)$ are subgroups of the group of all automorphisms of K, therefore

$$G(K, T) \subseteq G(K, F)$$

$\Rightarrow G(K, T)$ is a subgroup of $G(K, F)$.

Thus for any subfield T of K which contains F we have found a subgroup $G(K, T)$ of $G(K, F)$. Let us now define a mapping ψ from the set of subfields of K which contain F into the set of subgroups of $G(K, F)$ by the formula

$\psi(T)=G(K, T)$ for every subfield T of K which contains F.

ψ is onto.

For any subgroup H of $G(K, F)$, let K_H denote the fixed field of H *i.e.*, let $K_H=\{x \in K : \sigma(x)=x \ \forall \ \sigma \in H\}$.

Then K_H is a subfield of K. Also

$\sigma \in H \Rightarrow \sigma \in G(K, F)$ $\quad [\because H \subseteq G(K, F)]$

$\Rightarrow \sigma(\alpha)=\alpha \ \forall \ \alpha \in F$.

Thus if σ is any element of H, then $\sigma(\alpha)=\alpha \ \forall \ \alpha \in F$. Therefore $F \subseteq K_H$ and so K_H is a subfield of K containing F.

We have $\psi(K_H)=G(K, K_H)$ [by def. of ψ]
$=H$ [by theorem 36, since K is a finite extension of F and H is a subgroup of $G(K, F)$]

$\therefore$ ψ is onto.

ψ **is one-one.** Since K is a normal extension of a field F of characteristic 0, therefore K is the splitting field of some polynomial $f(x) \in F[x]$. If T is any subfield of K which contains F, then K is also the splitting field of $f(x)$ regarded as a polynomial over T. Therefore by theorem 38, K is a normal extension of T. Therefore, by the definition of normal extension the fixed field of $G(K, T)$ is T. Thus $K_{G(K, T)}=T$.

Now let T_1, T_2 be any two subfields of K which contain F. Then $\psi(T_1)=\psi(T_2)$

$\Rightarrow G(K, T_1)=G(K, T_2)$ [by def. of ψ]

$\Rightarrow$ the fixed field of $G(K, T_1)$=the fixed field of $G(K, T_2)$

$\Rightarrow K_{G(K, T_1)}=K_{G(K, T_2)}$

$\Rightarrow T_1=T_2$ [$\because K_{G(K, T)}=T$]

$\Rightarrow$ ψ is one-one.

Thus ψ gives us the desired one-to-one correspondence.

Now we shall prove the last three parts of the theorem.

(1) If T is any subfield of K containing F, then as proved above, K is a normal extension of T. Therefore by corollary to theorem 36, we have $o(G(K, T))=[K : T]$.

Since K is a normal extension of F, therefore

$$o(G(K, F))=[K : F]$$
$$=[K : T]\,[T : F] \quad \text{[by theorem 1]}$$
$$=o(G(K, T))\,[T : F].$$

$$\therefore \quad [T : F]=\frac{o(G(K, F))}{o(G(K, T))}=\text{index of } G(K, T) \text{ in } G(K, F).$$

(2) Let T be a subfield of K containing F. Then T is a normal extension of F

if and only if $\sigma(T) \subseteq T \;\forall\; \sigma \in G(K, F)$ [by theorem 39]

i.e., if and only if $\sigma(t) \in T \;\forall\; t \in T$ and $\forall\; \sigma \in G(K, F)$

i.e., if and only if $\tau[\sigma(t)]=\sigma(t) \;\forall\; t \in T, \forall\; \sigma \in G(K, F)$
and $\forall\; \tau \in G(K, T)$

[Note that $\tau \in G(K, T) \Leftrightarrow \tau(t)=t \;\forall\; t \in T$]

i.e., if and only if $\sigma^{-1}[\tau[\sigma(t)]]=\sigma^{-1}[\sigma(t)] \;\forall\; t \in T$,
$\forall\; \sigma \in G(K, F)$ and $\forall\; \tau \in G(K, T)$

(*a*) *Find the splitting field, K, of* x^3-2 *over* F_0 **(Meerut 1980,81)**

(*b*) *Prove that the Galois group of* x^3-2 *over* F_0 *is isomorphic to* S_3, *the symmetric group of degree* 3. **(Meerut 1980, 83, 84)**

(*c*) *Find the elements of the Galois group of* x^3-2 *over the field of rational numbers.* **(Punjab 1961; Aligarh 66)**

Solution. (*a*) F_0 is the field of rational numbers and x^3-2 is a polynomial over F_0. The roots of the polynomial x^3-2 in the field of complex numbers are $2^{1/3}$, $2^{1/3}\omega$, $2^{1/3}\omega^2$ where $2^{1/3}$ is the real cube root of 2 and ω is an imaginary cube root of unity. We claim that $K=F_0(2^{1/3}, \omega)$, is the splitting field of x^3-2 over F_0. We have $x^3-2=(x-2^{1/3})(x-2^{1/3}\omega)(x-2^{1/3}\omega^2)$.

Since $2^{1/3}$, $2^{1/3}\omega$, $2^{1/3}\omega^2$ are all in $F_0(2^{1/3},\omega)$, therefore $F_0(2^{1/3}, \omega)$ splits x^3-2 as a product of linear factors.

As proved in Ex. 1, page 512, the splitting field of x^3-2 over F_0 is of degree 6. If we know that $[F_0(2^{1/3},\omega):F_0]=6$, then $F_0(2^{1/3}, \omega)$ will be the splitting field of x^3-2 over F_0.

Now $[F_0(2^{1/3}, \omega):F_0]=[F_0(2^{1/3}, \omega):F_0(2^{1/3})]\,[F_0(2^{1/3}):F_0]$.

The minimal polynomial over F_0 belonging to $2^{1/3}$ is x^3-2. It is a polynomial of degree 3. Therefore $[F_0(2^{1/3}):F_0]=3$.

Since $\omega \notin F_0(2^{1/3})$, therefore ω cannot satisfy a polynomial of first degree over $F_0(2^{1/3})$. But ω satisfies the polynomial

$$x^2+x+1\in F_0(2^{1/3})[x]$$

Therefore x^2+x+1 is the minimal polynomial of ω over $F_0(2^{1/3})$. So $[F_0(2^{1/3}, \omega):F_0(2^{1/3})]=$ degree of $x^2+x+1=2$.

$$\therefore \quad [F_0(2^{1/3}, \omega):F_0]=2\times 3=6.$$

Hence $F_0(2^{1/3}, \omega)$ is the splitting field of x^3-2 over F_0.

(*b*) K is the splitting field of $x^3-2 \in F_0[x]$ over F_0. If $G(K, F_0)$ is the Galois group of x^3-2 over F_0, then

$$o(G(K:F_0))=[K:F_0]=6$$

The roots of x^3-2 in K are $\alpha=2^{1/3}$, $\beta=2^{1/3}\omega$, $\gamma=2^{1/3}\omega^2$.

The group $G(K, F_0)$ is isomorphic to a group of permutations on the set $\{\alpha, \beta, \gamma\}$.

Since $o(G(K, F_0))$ is 6, therefore the corresponding group of permutations on the set $\{\alpha, \beta, \gamma\}$ must also be of order 6. But S_3 is the only group of permtuations on the set $\{\alpha, \beta, \gamma\}$ which is of order 6. Hence $G(K, F_0)$ is isomorphic to S_3.

(*c*) Let $G(K, F_0)$ be the Galois group of x^3-2 over F_0. There are six distinct elements in the group $G(K, F_0)$ since $o(G(K, F_0))=6$. Let $S=\{\alpha, \beta, \gamma\}$ be the set consisting of the roots of x^3-2 in K. Let $\alpha=2^{1/3}$, $\beta=2^{1/3}\omega$, $\gamma=2^{1/3}\omega^2$. Let P be the set of those permutations on S which do not change those elements of S which are in F_0. There is a one-to-one correspondence between the elements of the set P and the elements of the group $G(K, F_0)$. Since S has no element which is in F_0, therefore there are six permutations on S which do not change those elements of S which are in F_0. These are I, $(\alpha\ \beta\ \gamma)$, $(\alpha\ \gamma\ \beta)$, $(\alpha\ \beta)$, $(\beta\ \gamma)$, $(\gamma\ \alpha)$. Thus these are

$$\begin{pmatrix}\alpha & \beta & \gamma\\ \alpha & \beta & \gamma\end{pmatrix}, \begin{pmatrix}\alpha & \beta & \gamma\\ \beta & \gamma & \alpha\end{pmatrix}, \begin{pmatrix}\alpha & \beta & \gamma\\ \gamma & \alpha & \beta\end{pmatrix},$$
$$\begin{pmatrix}\alpha & \beta & \gamma\\ \beta & \alpha & \gamma\end{pmatrix}, \begin{pmatrix}\alpha & \beta & \gamma\\ \alpha & \gamma & \beta\end{pmatrix}, \begin{pmatrix}\alpha & \beta & \gamma\\ \gamma & \beta & \alpha\end{pmatrix}.$$

If $\sigma\in G(K, F_0)$, then σ leaves every element of F_0 fixed. The field K is generated by α, β, γ over F_0. Further σ maps a root of x^3-2 onto a root of x^3-2. Therefore σ will be completely determined if we know under σ the images of any two of the three elements α, β, γ. If $\sigma_1, \sigma_2, \ldots, \sigma_6$ are the six elements of $G(K, F_0)$, then they will be determined by the following six sets of images :

$$\begin{pmatrix}\alpha & \beta\\ \alpha & \beta\end{pmatrix}, \begin{pmatrix}\alpha & \beta\\ \beta & \gamma\end{pmatrix}, \begin{pmatrix}\alpha & \beta\\ \gamma & \alpha\end{pmatrix}, \begin{pmatrix}\alpha & \beta\\ \beta & \alpha\end{pmatrix}, \begin{pmatrix}\alpha & \beta\\ \alpha & \gamma\end{pmatrix}, \begin{pmatrix}\alpha & \beta\\ \gamma & \beta\end{pmatrix}.$$

These can be written as

$$\begin{pmatrix}2^{1/3} & 2^{1/3}\omega\\ 2^{1/3} & 2^{1/3}\omega\end{pmatrix}, \begin{pmatrix}2^{1/3} & 2^{1/3}\omega\\ 2^{1/3}\omega & 2^{1/3}\omega^2\end{pmatrix}, \begin{pmatrix}2^{1/3} & 2^{1/3}\omega\\ 2^{1/3}\omega^2 & 2^{1/3}\end{pmatrix}$$
$$\begin{pmatrix}2^{1/3} & 2^{1/3}\omega\\ 2^{1/3}\omega & 2^{1/3}\end{pmatrix}, \begin{pmatrix}2^{1/3} & 2^{1/3}\omega\\ 2^{1/3} & 2^{1/3}\omega^2\end{pmatrix}, \begin{pmatrix}2^{1/3} & 2^{1/3}\omega\\ 2^{1/3}\omega^2 & 2^{1/3}\omega\end{pmatrix}.$$

Since $K=F_0\,(2^{1/3}, \omega)$, therefore $\sigma\in G(K, F_0)$ will be completely determined if we know the images of $2^{1/3}$ and ω under σ. Under each of the automorphisms $\sigma_1, \sigma_2, \ldots, \sigma_6$ we have already given the images of $2^{1/3}$. Now $\sigma_1(\omega)=\sigma_1\left(\dfrac{2^{1/3}\omega}{2^{1/3}}\right)=\dfrac{\sigma_1(2^{1/3}\omega)}{\sigma_1(2^{1/3})}=\dfrac{2^{1/3}\omega}{2^{1/3}}=\omega$,

$$\sigma_2(\omega)=\sigma_2\left(\frac{2^{1/3}\omega}{2^{1/3}}\right)=\frac{\sigma_2(2^{1/3}\omega)}{\sigma_2(2^{1/3})}=\frac{2^{1/3}\omega^2}{2^{1/3}\omega}=\omega,$$

$\sigma_3(\omega)=\omega$, $\sigma_4(\omega)=\omega^2=\sigma_5(\omega)=\sigma_6(\omega)$.

Therefore $\sigma_1, \sigma_2, \ldots, \sigma_6$ are given by

$$\begin{pmatrix}2^{1/3} & \omega\\ 2^{1/3} & \omega\end{pmatrix}, \begin{pmatrix}2^{1/3} & \omega\\ 2^{1/3}\omega & \omega\end{pmatrix}, \begin{pmatrix}2^{1/3} & \omega\\ 2^{1/3}\omega^2 & \omega\end{pmatrix},$$
$$\begin{pmatrix}2^{1/3} & \omega\\ 2^{1/3}\omega & \omega^2\end{pmatrix}, \begin{pmatrix}2^{1/3} & \omega\\ 2^{1/3} & \omega^2\end{pmatrix}, \begin{pmatrix}2^{1/3} & \omega\\ 2^{1/3}\omega^2 & \omega^2\end{pmatrix}.$$

§ 8. Construction with Straight-Edge (Ruler) and Compass.

Constructible number. Definition. *A real number α is said to be a constructible number if by the use of ruler and compass alone we can construct a line segment of length α. We assume that we are given some fundamental unit length.*

From our knowledge of high school geometry we know that with the help of ruler and compass alone we can draw a straight line perpendicular to a given line through a given point. Also we can draw a straight line parallel to a given line through a given point. From this we can easily show that if the real numbers α, β are constructible, then the real numbers $\alpha \pm \beta$, $\alpha\beta$ and α/β (if $\beta \neq 0$) are also constructible. Thus if W is the set of all constructible real numbers, then W is a subfield of the field of real numbers. Since $1 \in W$, therefore $p \in W$, where p is any integer. [Note that if p is a positive integer, then $p = 1+1+1+\ldots$upto p times. Further if $p = -q$ where q is a positive integer, then $p = -(1+1+\ldots$upto q times)]. Now if p and $0 \neq q$ are any two integers, then $\frac{p}{q}$ is constructible and so $\frac{p}{q} \in W$. Thus each rational number is constructible and so W contains F_0 the field of rational numbers.

Note. Throughout this section the words (*construct, constructible, construction*) will always mean by ruler and compass. Further the allowable steps of construction corresponding to the ruler and compass alone, are

(*a*) Drawing of a straight line through two points. [Use of ruler]

(*b*) Drawing of a circle with any point as centre and any line segment as radius. [Use of compass]

(*c*) Determination of a point or points as the intersection of (i) Two lines, (ii) A line and a circle, (iii) Two circles

Theorem 42. *Let F_0 be the field of rational numbers. The real number α is constructible if and only if we can find real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $\lambda_1^2 \in F_0$, $\lambda_2^2 \in F_0(\lambda_1)$, $\lambda_3^2 \in F_0(\lambda_1, \lambda_2), \ldots$, $\lambda_n^2 \in F_0(\lambda_1, \lambda_2, \ldots, \lambda_{n-1})$, and such that $\alpha \in F_0(\lambda_1, \lambda_2, \ldots, \lambda_n)$.*

Or

A real number α is constructible if and only if we can imbed α in a field obtained from the field of rational numbers by a finite number of quadratic extensions.

Proof. Suppose F is any subfield of the field of real numbers. The set of all those points (x, y) in the real Euclidean plane both of whose coordinates x and y are in F will be called *the plane F*. Any straight line joining two points in the plane of F has an equation of the form $ax+by+c=0$ where a, b, c are all in F Also the equation of any circle whose centre is a point in the plane of F and whose radius is an element of F is of the form

$$x^2+y^2+ax+by+c=0$$

where a, b, c are all in F. Such straight lines and circles are called *lines and circles in F.*

If two straight lines in the plane of F intersect in the real plane, then, the coordinates of their point of intersection are again in F because the process involves only the solution of two linear equations. On the other hand if a straight line and a circle in the plane of F intersect in the real plane, then the determination of the coordinates of their points of intersection leads to the solution of a quadratic equation in one unknown with coefficients belonging to F. Thus the coordinates of these points of intersection either belong to F or to a quadratic extension of the form $F(\sqrt{\gamma})$ where γ is a positive member of F but is not the square of a member of F. [It should be noted that if γ is a positive member of F, then $F(\sqrt{\gamma})=F$ if γ is the square of a member of F and $F(\sqrt{\gamma})$ is a proper extension of F if γ is not the square of a member of F In the former case $[F(\sqrt{\gamma}):F]=1$ and in the later case $[F(\sqrt{\gamma}):F]=2$].

Finally, the intersection of two circles in F can be realized as that of a line in F and a circle in F. For example if two circles in F are $x^2+y^2+a_1x+b_1y+c_1=0$ and $x^2+y^2+a_2x+b_2y+c_2=0$, their intersection is the intersection of either of these with the line $(a_1-a_2)x+(b_1-b_2)y+(c_1-c_2)=0$ which passes through their common points. Thus if two circles in the plane of F intersect in the real plane, then their point of intersection is either in the plane of F or in the plane of $F(\sqrt{\gamma})$ for some positive γ in F.

From the above discussion we conclude that lines and circles of F lead us to points either in F or in quadratic extensions of F. Suppose in place of being in F, we are now in $F(\sqrt{\gamma_1})$ for some quadratic extension of F. Then lines and circles in $F(\sqrt{\gamma_1})$ intersect in points in the plane of $F(\sqrt{\gamma_1}, \sqrt{\gamma_2})$ where γ_2 is a positive member in $F(\sqrt{\gamma_1})$. Thus a point is constructible from F if we can find real numbers $\lambda_1, \ldots, \lambda_n$ such that $\lambda_1^2 \in F$, $\lambda_2^2 \in F(\lambda_1)$, $\lambda_3^2 \in F(\lambda_1, \lambda_2), \ldots, \lambda_n^2 \in F(\lambda_1, \ldots, \lambda_{n-1})$, and such that the point is in the plane of $F(\lambda_1, \ldots, \lambda_n)$. Also from our knowledge of high school geometry we know that if a positive real number γ is constructible, then $\sqrt{\gamma}$ is also constructible. Therefore if $\lambda_1, \ldots, \lambda_n$ are real numbers such that $\lambda_1^2 \in F$, $\lambda_2^2 \in F(\lambda_1)$, $\lambda_3^2 \in F(\lambda_1, \lambda_2)$, $\ldots$, $\lambda_n^2 \in F(\lambda_1, \ldots, \lambda_{n-1})$, then a point in the plane of $F(\lambda_1, \ldots, \lambda_n)$ is constructible from F. Hence a point is constructible from F if and only if we can find a finite number of real numbers $\lambda_1, \ldots, \lambda_n$ such that

(1) $[F(\lambda_1) : F] = 1$ or 2,

(2) $[F(\lambda_1, \ldots, \lambda_i) : F(\lambda_1, \ldots, \lambda_{i-1})] = 1$ or 2 for $i = 1, 2, \ldots, n$ and such that our point lies in the plane of $F(\lambda_1, \ldots, \lambda_n)$.

Now we know that every number belonging to the field F_0 of rational numbers is constructible. Therefore in the above discussion if we take the field F_0 in place of the field F, then we conclude that a real number α is constructible if and only if we can imbed α in a field obtained from F_0 by a finite number of quadratic extensions of F_0.

This completes the proof of the theorem.

Theorem 43. *If a real number α is constructible, then α lies in some finite extension K of the field of rational numbers of degree which is some power of 2.* **(Delhi 1970; Dibrugarh 67)**

Proof. Let F_0 be the field of rational numbers. We know that if a real number α is constructible, then α can be imbedded in a field obtained from F_0 by a finite number of quadratic extensions. Thus if a real number α is constructible, then we can find real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that

$$\lambda_1^2 \in F_0,\ \lambda_2^2 \in F_0(\lambda_1),\ \lambda_3^2 \in F_0(\lambda_1, \lambda_2), \ldots,$$

$\lambda_n^2 \in F_0(\lambda_1, \lambda_2, \ldots, \lambda_{n-1})$, and such that $\alpha \in F_0(\lambda_1, \ldots, \lambda_n)$.

We have $\quad [F_0(\lambda_1) : F_0] = 1 \quad$ or $\quad 2$.

Also $\quad [F_0(\lambda_1, \ldots, \lambda_i) : F_0(\lambda_1, \ldots, \lambda_{i-1})] = 1 \quad$ or $\quad 2$

for $i = 1, 2, \ldots, n$.

Now $\alpha \in F_0(\lambda_1, \ldots, \lambda_n)$. We have

$$[F_0(\lambda_1, \ldots, \lambda_n) : F_0]$$
$$= [F_0(\lambda_1, \ldots, \lambda_n) : F_0(\lambda_1, \ldots, \lambda_{n-1})]\, [F_0(\lambda_1, \ldots, \lambda_{n-1}) : F_0(\lambda_1, \ldots, \lambda_{n-2})] \cdots [F_0(\lambda_1, \lambda_2) : F_0(\lambda_1)]\, [F_0(\lambda_1) : F_0].$$

Since each term in the product is either 1 or 2, we get that

$$[F_0(\lambda_1, \ldots, \lambda_n) : F_0] = 2^r$$

where r is some non-negative integer.

Thus $\alpha \in F_0(\lambda_1, ..., \lambda_n)$ which is a finite extension of F_0 and has degree a power of 2.

Theorem 44. *If the real number α satisfies an irreducible polynomial over the field of rational numbers of degree k, and if k is not a power of 2, then α is not constructible.*

Proof. Let F_0 be the field of rational numbers. Suppose α is a constructible real number. Then there is a finite extension K of F_0 such that $\alpha \in K$ and such that $[K : F_0] = 2^r$ where r is some non-negative integer. Since $F_0(\alpha)$ is a subfield of K containing F_0, therefore by theorem 2, $[F_0(\alpha) : F_0]$ is a divisor o $[K : F_0] = 2^r$. Therefore $[F_0(\alpha) : F_0]$ must also be some power of 2. Thus we conclude that if α is constructible, then $[F_0(\alpha) : F_0]$ must be some power of 2. Now it is given that α satisfies some irreducible polynomial of degree k over F_0. Therefore

$$[F_0(\alpha) : F_0] = k.$$

Since k is given to be not a power of 2, therefore α cannot be constructible. This proves the theorem.

Note. The above theorem gives us a very important criterion of non-constructibility.

Solved Examples

Ex. 1. *Show that the polynomial* $8x^3 - 6x - 1$ *is irreducible over the field of rational numbers.*

Solution. Let $f(x) = 8x^3 - 6x - 1$. Then

$$f(x-1) = 8(x-1)^3 - 6(x-1) - 1 = 8x^3 - 24x^2 + 18x - 3.$$

Obviously $f(x)$ is irreducible over the field of rational numbers if and only if $f(x-1)$ is irreducible over the field of rational numbers.

Now $f(x-1) = 8x^3 - 24x^2 + 18x - 3$ is a polynomial with integer coefficients. Also 3 is a prime number such that 3 divides each of the coefficients of $f(x-1)$ except the coefficient of x^3 which is 8. Also 3^2 is not a divisor of the constant term which is -3. Therefore by Eisenstein's criterion of irreducibility $f(x-1)$ is irreducible over the field of rational numbers. Hence $f(x)$ is irreducible over the field of rational numbers.

Ex. 2. *Show that it is impossible, by straight edge and compass alone, to trisect* 60°. **(Meerut 1978, 80, 81, 82, 84; Kanpur 80)**

Solution. Suppose it is possible to trisect 60° by stright-edge and compass alone. Then the length $\alpha = \cos 20°$ would be constructible. We have the identity

$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta.$$

Putting $\theta = 20°$ in this identity, we get

$$\cos 60° = 4\cos^3 20° - 3\cos 20°$$

or $$\tfrac{1}{2} = 4\alpha^3 - 3\alpha \qquad [\because \cos 60° = \tfrac{1}{2}, \cos 20° = \alpha]$$

or $$8\alpha^3 - 6\alpha - 1 = 0.$$

Thus the real number α is a root of the polynomial $8x^3 - 6x - 1$ over the field of rational numbers. Now $8x^3 - 6x - 1$ is an irreducible polynomial over the field of rational numbers. The degree of this polynomial is 3 which is definitely not a power of 2. Therefore by theorem 44, α is not constructible. Hence our initial assumption is wrong. Thus it is not possible to trisect 60° with straight edge and compass alone.

Ex. 3. *Show that it is impossible by straight-edge and compass alone to duplicate the cube.* (Meerut 1979, 81, 83; Dibrugarh 67)

Solution. We are to show that it is not possible by straight edge and compass alone to construct a cube whose volume is twice that of a given cube. Without loss of generality, let us suppose that the given cube is the unit cube. Then we are to show that it is not possible to construct a length α such that

$$\alpha^3 = 2 \quad \text{or} \quad \alpha^3 - 2 = 0.$$

Now $x^3 - 2$ is an irreducible polynomial over the field of rational numbers. The degree of this polynomial is 3 which is certainly not a power of 2. Therefore, by theorem 44, if the real number α is a root of this polynomial, then α is not constructible.

Ex. 4. *Show that it is impossible to construct a regular septagon by straight edge and compass alone.* (Meerut 1978, 79, 84 P)

Solution. The sum of the exterior angles of a regular polygon is equal to 2π. Therefore each exterior angle of a regular septagon is equal to $2\pi/7$. If it is possible to construct a regular septagon by straight edge and compass alone, then the length $\alpha = 2\cos 2\pi/7$ is constructible.

Let $\theta = 2\pi/7$. Then $7\theta = 2\pi$ or $4\theta = 2\pi - 3\theta$.

$$\therefore \quad \sin 4\theta = \sin(2\pi - 3\theta) = -\sin 3\theta$$

or $$2.2\sin\theta\cos\theta\,(2\cos^2\theta - 1) + (3\sin\theta - 4\sin^3\theta) = 0$$

or $$\sin\theta\,[8\cos^3\theta - 4\cos\theta + 3 - 4(1 - \cos^2\theta)] = 0$$

or $$8\cos^3\theta + 4\cos^2\theta - 4\cos\theta - 1 = 0$$

$$\left[\because \sin\theta \neq 0 \text{ if } \theta = \frac{2\pi}{7}\right]$$

or $$(2\cos\theta)^3 + (2\cos\theta)^2 - 2\,(2\cos\theta) - 1 = 0.$$

Therefore $2\cos(2\pi/7)$ is a root of the equation $x^3+x^2-2x-1=0$. Thus $\alpha=2\cos 2\pi/7$ satisfies the polynomial x^3+x^2-2x-1. But this polynomial is irreducible over the field of rational numbers. Its degree is 3 which is not a power of 2. Therefore α is not constructible. Hence it is not possible to construct a regular septagon by straight-edge and compass alone.

Ex. 5. *Show that the regular pentagon is constructible.*

Solution. A regular pentagon will be constructible if it is possible to construct a length $\alpha=\cos 2\pi/5$. We have

$$\cos\frac{2\pi}{5}=\cos\frac{2}{5}(180^\circ)=\cos 72^\circ=\sin 18^\circ=\frac{-1+\sqrt{5}}{4}.$$

Since $\sqrt{5}=\sqrt{(1^2+2^2)}$ is constructible, therefore $\alpha=\dfrac{-1+\sqrt{5}}{4}$ is constructible. Hence it is possible to construct the regular pentagon.

Ex. 6. *Show that the regular hexagon is constructible.*

Solution. A regular hexagon will be constructible if it is possible to construct a length $\alpha=\cos\dfrac{2\pi}{6}=\cos\dfrac{\pi}{3}=\dfrac{1}{2}$. Now $\frac{1}{2}$ is constructible. Therefore the regular hexagon is constructible.

Ex. 7. *Show that it is not possible to construct a square whose area is equal to that of a given circle*

Solution. Without loss of generality we can assume that the radius of the given circle is unity. Then the area of the given circle is π. So we are to construct a length α such that

$$\alpha^2=\pi \quad \text{or} \quad \alpha^2-\pi=0$$

Since π is transcendental, therefore it cannot be the root of an equation whose coefficients belong to the field F_0 of rational numbers. Thus $F_0(\pi)$ is not algebraic over F_0 so that $F_0(\pi)$ is of infinite degree over F_0. Thus π and so also $\sqrt{\pi}$ are not constructible.

§ 9. Solvability by Radicals Consider the polynomial x^2+3x+4 over the field of rational numbers F_0. Using the quadratic formula for its roots, we see that its roots are $\frac{1}{2}[-3\pm\sqrt{(-7)}]$. Thus the field $F=F_0(\sqrt{7}i)$ is the splitting field of x^2+3x+4 over F_0. Consequently we may say that there is an element $\gamma=-7$ in F_0 such that the extension field $F_0(\omega)$ where $\omega^2=\gamma$ is such that it contains all the roots of x^2+3x+4.

Definition. *Suppose F is a field and $p(x)$ is a polynomial in $F[x]$. Then $p(x)$ is said to be solvable by radicals over F if it is possible to determine a finite sequence of fields*

$$F_1=F(\omega_1),\ F_2=F_1(\omega_2)=F(\omega_1,\omega_2),\ \ldots,\ F_k=F_{k-1}(\omega_k)$$
$$=F(\omega_1,\ldots,\omega_k)$$

such that $\omega_1^{r_1}\in F,\ \omega_2^{r_2}\in F_1,\ \ldots,\ \omega_k^{r_k}\in F_{k-1}$, *such that the roots of* $p(x)$ *all lie in* F_k.

It is obvious that

$$F\subseteq F_1\subseteq F_2\subseteq\cdots\subseteq F_k.$$

Remark If K is the splitting field of $p(x)$ over F, then $p(x)$ is solvable by radicals over F if it is possible to find a sequence of fields as in the above definition such that $K\subseteq F_k$.

Also it can be proved that if $p(x)$ is solvable by radicals over F, then we can find a sequence of fields

$$F\subseteq F_1=F(\omega_1)\subseteq F_2=F_1(\omega_2)\subseteq\cdots\subseteq F_k=F_{k-1}(\omega_k)$$

where $\omega_1^{r_1}\in F,\ \omega_2^{r_2}\in F_1,\ldots,\ \omega_k^{r_k}\in F_{k-1}$, F_k containing all the roots of $p(x)$ such that F_k is *normal over F.*

From the classical theory of algebraic equations, we know that the equations of second, third and fourth degrees over the field of rational numbers can be solved by radicals. If $p(x)$ is a given polynomial over a field F, then we want to know whether it is solvable by radicals or not. For this purpose we shall give a criterion for the solvability by radicals of a polynomial $p(x)$ over a field F in terms of the Galois group of that polynomial. But first we need a result about the Galois group of a certain type of polynomial.

Theorem 45. *Suppose that the field F contains all n^{th} roots of unity (for some particular n) and suppose that a is a non-zero element of F. Let $x^n-a\in F[x]$ and let K be its splitting field over F. Then*

(1) *$K=F(u)$ where u is any root of x^n-a.*

(2) *The Galois group of x^n-a over F is abelian.*

Proof (1) We know that the nth roots of unity are given by $\cos(2r\pi/n)+i\sin(2r\pi/n)=e^{2r\pi i/n}$, where $r=0,1,\ldots,n-1$.

It is given that F contains all nth roots of unity. Therefore, in particular, F contains $\alpha=e^{2\pi i/n}$. Note that $\alpha^n=1$ but $\alpha^m\neq 1$ for $0<m<n$.

If $u\in K$ is any root of x^n-a, then we claim that $u, \alpha u, \alpha^2 u, \ldots, \alpha^{n-1}u$ are all the roots of x^n-a. That they are roots is obvious. Also no two of these are equal because

$$\alpha^i u=\alpha^j u \qquad \text{where } 0\leq i<j<n$$

$\Rightarrow (\alpha^i-\alpha^j)\, u=0$

$\Rightarrow \alpha^i-\alpha^j=0$ [$\because$ u cannot be 0]

$\Rightarrow \alpha^i=\alpha^j$

$\Rightarrow \alpha^{j-i}=1$ where $0<j-i<n$.

But this is impossible since $\alpha^m \neq 1$ for $0<m<n$. Therefore no two of the roots $u, \alpha u, \ldots, \alpha^{n-1}u$ are equal. Thus $u, \alpha u, \alpha^2 u, \ldots, \alpha^{n-1}u$ are n distinct roots of x^n-a which cannot have more than n roots. Therefore $u, \alpha u, \alpha^2 u, \ldots, \alpha^{n-1}u$ are all the roots of x^n-a. Since $\alpha \in F$, therefore all of $u, \alpha u, \ldots, \alpha^{n-1}u$ are in $F(u)$. Thus $F(u)$ splits x^n-a. Further $F(u)$ is the smallest field containing u and F. Therefore no proper subfield of $F(u)$ which contains F also contains u. Hence no proper subfield of $F(u)$ can split x^n-a. It follows that $F(u)$ is the splitting field of x^n-a and thus

$$K=F(u).$$

(2) Let $G(K, F)$ be the Galois group of

$$f(x)=x^n-a \in F[x].$$

Let σ, τ be any two elements of $G(K, F)$. Then σ, τ are automorphisms of $K=F(u)$ leaving every element of F fixed. Since $u \in K$ is a root of $f(x)$, therefore both $\sigma(u)$ and $\tau(u)$ are roots of $f(x)$. [See Ex. 5 page 514]. Then we must have $\sigma(u)=\alpha^i u$ and $\tau(u)=\alpha^j u$ for some i and j.

Now $\sigma\tau(u)=\sigma(\tau(u))=\sigma(\alpha^j u)=\sigma(\alpha^j)\,\sigma(u)$

$=\alpha^j \sigma(u)$ [$\because$ $\alpha^j \in F \Rightarrow \sigma(\alpha^j)=\alpha^j$]

$=\alpha^j \alpha^i u=\alpha^{i+j}u.$

Similarly $\tau\sigma(u)=\alpha^{i+j}u.$

Thus $\sigma\tau$ and $\tau\sigma$ agree on u and on F and consequently they agree on all of $K=F(u)$. It follows that $\sigma\tau=\tau\sigma$ and so the Galois group $G(K, F)$ is abelian.

Remark. The above theorem implies that if F has all nth roots of unity, then by adjoining one root of x^n-a to F where $a \in F$, we obtain the whole splitting field of x^n-a and accordingly this must be a normal extension of F.

Before going through the remaining part of this section students are advised to revise the concept of solvable groups given in § 13 of chapter 3 on Groups (continued).

Theorem 46 *Let F be a field of characteristic 0 containing all nth roots of unity for every integer n. If $p(x) \in F[x]$ is solvable by radicals over F, then the Galois group over F of $p(x)$ is a solvable group.*

Proof. Let K be the splitting field of $p(x)$ over F and let $G(K, F)$ be the Galois group of $p(x)$ over F. It is given that $p(x)$ is solvable by radicals over F. Therefore by the definition of solvability of a polynomial by radicals, there exists a sequence of fields

$$F \subseteq F_1 = F(\omega_1) \subseteq F_2 = F_1(\omega_2) \subseteq \dots \subseteq F_k = F_{k-1}(\omega_k)$$

where $\omega_1^{r_1} \in F$, $\omega_2^{r_2} \in F_1, \dots, \omega_k^{r_k} \in F_{k-1}$ and where $K \subseteq F_k$. Without any loss of generality we may regard F_k as a normal extension of F. [See our remark on page 553]. Since F_k is a normal extension of F, therefore F_k is also a normal extension of any intermediate field. Thus F_k is a normal extension of each F_i.

By our remark after theorem 45, each F_i is a normal extension of F_{i-1}. Also since F_k is a normal extension of F_{i-1}, therefore by the fundamental theorem of Galois theory it follows that $G(F_k, F_i)$ is a normal subgroup of $G(F_k, F_{i-1})$. Now consider the chain

$$G(F_k, F) \supseteq G(F_k, F_1) \supseteq G(F_k, F_2) \supseteq \dots \supseteq$$
$$G(F_k, F_{k-1}) \supseteq (e). \qquad \dots(1)$$

We have just mentioned that each subgroup in this chain is a normal subgroup of the preceding one. Since F_i is a normal extension of F_{i-1}, therefore by the fundamental theorem of Galois theory the Galois group $G(F_i, F_{i-1})$ of F_i over F_{i-1} is isomorphic to the quotient group $G(F_k, F_{i-1})/G(F_k, F_i)$. But by theorem 45, $G(F_i, F_{i-1})$ is an abelian group and the isomorphic image of an abelian group is also an abelian group. Therefore each quotient group $G(F_k, F_{i-1})/G(F_k, F_i)$ of the chain (i) is abelian. Hence by definition of a solvable group, the group $G(F_k, F)$ is solvable.

Now K is the splitting field of $p(x)$ over F and so K is a normal extension of F. Also $K \subseteq F_k$. Therefore by the fundamental theorem of Galois theory $G(F_k, K)$ is a normal subgroup of $G(F_k, F)$ and $G(K, F)$ is isomorphic to $G(F_k, F)/G(F_k, K)$. It follows that $G(K, F)$ is a homomorphic image of $G(F_k, F)$ which is a solvable group. But every homomorphic image of a solvable group is also a solvable group. [See corollary to theorem 3 page 242]. Therefore the group $G(K, F)$ is solvable. Since $G(K, F)$ is the Galois group of $p(x)$ over F, therefore the proof of the theorem is complete.

Note 1. The converse of the above theorem is also true :

If the Galois group of $p(x)$ over F is solvable, then $p(x)$ is solvable by radicals over F.

Note 2. Theorem [illegible] and its converse are true even if F does not contain roots of unity.

In the end of this section we shall give the famous theorem of Abel in connection with the solvability by radicals. First we shall explain the meaning of *the general polynomial of degree n over a field.*

Let $F(a_1, \ldots, a_n)$ be the field of rational functions, in the n variables $a_1, \ldots, a_n$ over F. The particular polynomial $p(x)=x^n+a_1x^{n-1}+\ldots+a_n$ over the field $F(a_1, \ldots, a_n)$ is called the general polynomial of degree n over the field F. We say that it is solvable by radicals if it is solvable by radicals over the field $F(a_1, \ldots, a_n)$. It can be easily shown that the Galois group of the polynomial $p(x)=x^n+a_1x^{n-1}+\ldots+a_n$ over $F(a_1, \ldots, a_n)$ is P_n, the symmetric group of degree n.

Theorem 47. (Abel's theorem). *The general polynomial of degree $n \geqslant 5$ is not solvable by radicals.* **(I.A.S. 1975)**

Proof. Let $F(a_1, \ldots, a_n)$ be the field of rational functions in the n variables $a_1, \ldots, a_n$. Then the Galois group of the polynomial $p(x)=x^n+a_1x^{n-1}+\ldots+a_n$ over $F(a_1, \ldots, a_n)$ is P_n, where P_n is the symmetric group of degree n. By theorem 5 page 244, P_n is not a solvable group when $n \geqslant 5$. Therefore by theorem 46, the polynomial $p(x)$ is not solvable by radicals over $F(a_1, \ldots, a_n)$ when $n \geqslant 5$.

§ 10. Finite Fields.

Finite field. Definition. *A field having only a finite number of elements is called a finite field.* If p is a prime number, then the ring I_p of integers modulo p is a field. This field has p elements and so it is an example of a finite field. We know that the characteristic of a field is either zero or a prime number. Suppose F is a finite field having n elements. Since F is a group of order n with respect to addition, therefore by a corollary to Lagrange's theorem we have $1+1+1+\ldots$ upto n times$=0$. So the characteristic of F cannot be zero. Hence the characteristic of a finite field is always a prime number.

Theorem 48. *Let F be a finite field with q elements and suppose that $F \subseteq K$ where K is also a finite field. Then K has q^n elements where $n=[K:F]$.* **(Meerut 1978)**

Proof. Since F is a sub-field of the field K, therefore K can be regarded as a vector space over the field F. The number of elements in the set K is finite. So the vector space $K(F)$ is definitely finite dimensional. Let the dimension of the vector space $K(F)$ be n. Then $[K : F]=n$. Let $\{\beta_1, \beta_2, \ldots, \beta_n\}$ be a basis of the vector space $K(F)$. Then every element in K can be uniquely expressed in the form $a_1\beta_1+a_2\beta_2+\ldots+a_n\beta_n$ where $a_1, a_2, \ldots, a_n$ are all in F. Then the number of elements in K is the number of $a_1\beta_1+a_2\beta_2+\ldots+a_n\beta_n$ as the $a_1, a_2, \ldots, a_n$ range over F. Since F has q elements, therefore each of the n coefficients a's can have q values. So K must have q^n elements.

Theorem 49. *Let F be a finite field. Then F has p^m elements where the prime number p is the characteristic of F.*

(Meerut 1972; Nagarjuna 79, 80; Banaras 71)

Proof. Since F is a finite field, therefore the characteristic of F is a prime number, say, p. Now F is a field of characteristic p. Therefore F contains a sub-field F_0 isomorphic to the field $\mathbf{I}_p$ of integers modulo p. The field $\mathbf{I}_p$ has p elements. Therefore the field F_0 has also p elements. Now F is a finite field and F_0 is a subfield of F. Let $[F : F_0]=m$. Then, by theorem 48, the number of elements in $F=p^m$.

Theorem 50. *If the finite field F has $q=p^m$ elements, then every $a \in F$ satisfies the relation $a^q=a$* **(Meerut 1981, 82P)**

Proof. Let $a \in F$ and $a=0$. We have $0^q=0$. Therefore $a=0$ satisfies the relation $a^q=a$.

Now let $a \in F$ and $a \neq 0$. The non-zero elements of the field F form a group with respect to multiplication. The identity element of this group is 1 and the order of this group is $q-1$. Therefore by a corollary to Lagrange's theorem, we have

$$a^{q-1}=1 \text{ for all } a \neq 0 \text{ in } F. \qquad \ldots(1)$$

Multiplying both sides of (1) by a, we get $a^q=a$.

Thus we have $a^q=a \ \forall \ a \in F$.

Corollary. *If the finite field F has $q=p^m$ elements, then the polynomial x^q-x in $F[x]$ factors in $F[x]$ as*

$$x^q-x=\prod_{\lambda \in F} (x-\lambda).$$

(Meerut 1980, 82P, Banaras 71)

Proof. As proved in theorem 50, we have $a^q=a \ \forall \ a \in F$. Therefore every a in F satisfies the polynomial $x^q-x \in F[x]$.

Hence by factor theorem $x-\lambda$ is a divisor of x^q-x for all $\lambda \in F$. Let $a_1, a_2, \ldots, a_q$ be the q elements of the field F. Then the product $(x-a_1)(x-a_2)\ldots(x-a_q)$ is a divisor of x^q-x, being a product of relatively prime polynomials each dividing x^q-x. Now $(x-a_1)(x-a_2)\ldots(x-a_q)$ is a monic polynomial of degree q and it is a divisor of the monic polynomial x^q-x of degree q. Therefore we have

$$x^q-x=(x-a_1)(x-a_2)\ldots(x-a_q)=\prod_{\lambda\in F}(x-\lambda).$$

Theorem 51 *Any two finite fields having the same number of elements are isomorphic.* (Meerut 1972 ; Nagarjuna 79, 80)

Proof. Let F be a finite field of characteristic p. Let the number of elements in F be equal to $q=p^n$. First we shall show that F can be regarded as the splitting field of the polynomial $x^q-x\in \mathbf{I}_p[x]$ where $\mathbf{I}_p$ is the field of integers modulo p.

Since F is a field of finite characteristic p, therefore it contains a subfield F_0 isomorphic to the field $\mathbf{I}_p$. So, without any loss of generality, we can regard F as an extension of the field $\mathbf{I}_p$. Since the field F has q elements, therefore by theorem 50, we have $a^q=a \ \forall \ a\in F$. So every a in F satisfies the polynomial $x^q-x \in \mathbf{I}_p[x]$. Since the degree of the polynomial x^q-x is q, therefore it can have at most q roots in any extension field of $\mathbf{I}_p$. But we have just shown that all the q elements of the field F are the roots of x^q-x. Therefore we have $x^q-x=\prod_{\lambda\in F}(x-\lambda)$. Thus the polynomial $x^q-x \in \mathbf{I}_p[x]$ splits in the field F. But this polynomial cannot split in any smaller field for that field would have to have all the roots of this polynomial and so would have to have at least q elements. Note that the roots of this polynomial are all distinct. [See theorem 27 on page 520]. Consequently F is the splitting field of $x^q-x\in \mathbf{I}_p[x]$.

Now suppose that F^* is any other finite field having $q=p^n$ elements. Then as shown above F^* is also the splitting field of $x^q-x\in \mathbf{I}_p[x]$. But any two splitting fields of $x^q-x \in \mathbf{I}_p[x]$ must be isomorphic. [See corollary to theorem 22 on page 512]

Hence $F^*\cong F$.

Theorem 52. *For every prime number p and every positive integer m there exists a field having p^m elements.*

(Banaras 1972 ; Meerut 77, 80, 81)

Proof. Since p is a prime number, therefore I_p, the ring of integers modulo p, is a finite field of characteristic p. Consider the polynomial $x^{p^m}-x \in I_p[x]$. Let K be the splitting field of this polynomial. In K let $F=\{a \in K : a^{p^m}=a\}$. Then F is a subset of K and it consists of those elements of K which are the roots of $x^{p^m}-x$. But by theorem 27 the roots of the polynomial $x^{p^m}-x$ are all distinct and so this polynomial has p^m distinct roots in the field K. Therefore the set F has p^m elements. We now claim that F is a subfield of K. Let $a, b \in F$. Then $a^{p^m}=a$ and $b^{p^m}=b$. Since the characteristic of the field K is p, therefore $(a \pm b)^{p^m}=a^{p^m} \pm b^{p^m}=a \pm b$. Thus $a, b \in F \Rightarrow a \pm b \in F$. Also $(ab)^{p^m}=a^{p^m}\, b^{p^m}=ab$. Therefore $a, b \in F \Rightarrow ab \in F$. Similarly we can show that if $a \in F$ and $0 \neq b \in F$, then $a/b \in F$. Consequently F is a subfield of K and so F itself is a field and it has p^m elements. Thus we have shown that for every prime number p and every positive integer m there exists a field F having p^m elements.

Theorem 53. *For every prime number p and every positive integer m there is a unique field having p^m elements.*

(Banaras 1972 ; Meerut 76, 83 ; Andhra 77)

Proof. As shown in theorem 52, for every prime number p and every positive integer m there exists a finite field having p^m elements. [Give the proof here]. Also by theorem 51, any two finite fields having the same number of elements are isomorphic. Hence for every prime number p and every positive integer m there exists a unique field having p^m elements. This field is sometimes called the *Galois field* $GF\ [p^m]$. **(Andhra 1977)**

Definition *Let ω be an nth root of unity. If s is the smallest positive integer such that $\omega^s=1$, we shall say that s is the* **order** *of ω. If the order of ω is n we shall call ω a* **primitive** *nth root of unity.*

The following theorem will completely describe the structure of the multiplicative group of the Galois field $GF\ [p^m]$. In the proof of this theorem we shall use one result about abelian groups which we give in the form of the following Lemma.

Lemma. *In an abelian group the product $c_1c_2\ldots c_r$ of elements*

c_i whose orders are powers $p_i^{e_i}$ of distinct primes has order exactly

$$p_1^{e_1}p_2^{e_2}\ldots p_r^{e_r}=h.$$

The easy proof of this lemma has been left as an exercise for the reader. Show that the order divides h, but fails to divide h/p_i for any i.

Theorem 54. *The multiplicative group of non-zero elements of a finite field is cyclic.* **(Banaras 1972)**

Proof. Let F be a finite field having q elements. The non-zero elements of F form a group of order $q-1$ with respect to multiplication. The identity element of this group is 1. Therefore we have $a^{q-1}=1$ for all $a\neq 0$ in F. Thus each non-zero element in F is a $(q-1)$th root of unity, in the sense that it satisfies the equation $x^{q-1}=1$. Let F_0 denote the set of non-zero elements of F. If we find an element $a\in F_0$ such that the order of a as an element of the multiplicative group F_0 is $q-1$, then F_0 will turn out to be a cyclic group. [See theorem 5 on page 173 in the chapter 2 of Groups]. To achieve this aim let us write $q-1$ as product of powers of distinct primes

$$q-1=p_1^{e_1}p_2^{e_2}\ldots p_r^{e_r}\ (0<p_1<p_2<\ldots<p_r).$$

For each i, $p_i^{e_i}$ is a divisor of $q-1$. Therefore the roots of $x^P=1$, where $P=p_i^{e_i}$, are all roots of $x^{q-1}=1$, hence all lie in F. Let us put $Q=p_i^{e_i-1}$. Then of all the $P=p_i^{e_i}$ distinct roots of the equation $x^P=1$, exactly $Q=p_i^{e_i-1}$ satisfy the equation $x^Q=1$. Therefore F contains at least one root $c=c_i$ of $x^P=1$ which does not satisfy $x^Q=1$. Then $c_i\in F$ is such that $c_i^P=1$ but $c_i^Q\neq 1$. Therefore the element c_i has order $P=p_i^{e_i}$ in the multiplicative group of non-zero elements of F. By the above Lemma the product $c_1c_2\ldots c_r$ is an element of order $p_1^{e_1}p_2^{e_2}\ldots p_r^{e_r}=q-1$. Thus $c_1c_2\ldots c_r$ is the required element of order $q-1$ of the multiplicative group of non-zero elements of F. Hence the multiplicative group of non-zero elements of F is cyclic.

Theorem 55. *Every finite field of characteristic p has an automorphism $a\leftrightarrow a^p$.* **(Meerut 1981)**

Proof. Let F be a finite field of characteristic p. Let $\phi : F\to F$ such that $\phi(a)=a^p\ \forall\ a\in F$.

Offset Printers-Vasudha Printers, Delhi Ph. : 266563

ϕ is one-one. Let $a, b \in F$. Then

$\phi(a)=\phi(b) \Rightarrow a^p=b^p \Rightarrow a^p-b^p=0$

$\Rightarrow (a-b)^p=0$ [Note that in a field of characteristic p, we have $(a-b)^p=a^p-b^p$]

$\Rightarrow a-b=0 \Rightarrow a=b \Rightarrow \phi$ is 1−1.

ϕ is onto. Since the set F is finite, therefore ϕ is one-one $\Rightarrow \phi$ is also onto.

ϕ preserves addition and multiplication. Let $a, b \in F$. Then

$\phi(a+b)=(a+b)^p$

$=a^p+b^p$ [Note that in a field of characteristic p, we have $(a+b)^p=a^p+b^p$]

$=\phi(a)+\phi(b)$.

Also $\phi(ab)=(ab)^p=a^pb^p=\phi(a)\,\phi(b)$.

Hence ϕ is an automorphism of the field F.

Corollary *In a finite field of characteristic p, every element has a p^{th} root.*

Theorem 56. *If F is a finite field and $\alpha \neq 0$, $\beta \neq 0$ are two elements of F then we can find elements a and b in F such that*

$1+\alpha a^2+\beta b^2=0$. **(Banaras 1971 ; Meerut 80, 82)**

Proof. Case 1. Suppose the characteristic of F is 2. Then F has 2^m elements and every element x in F satisfies $x^{2^m}=x$. [See theorem 50]. Thus every element in F is a square of some element in F. In particular $\alpha^{-1}=c^2$ for some $c \in F$. Taking $a=c$ and $b=0$, we have

$$1+\alpha a^2+\beta b^2=1+\alpha c^2+\beta 0=1+\alpha\alpha^{-1}+0=1+1=0.$$

Note that if the characteristic of F is 2, then $1+1=0$.

Case 2. Suppose the characteristic of F is a prime number p other than 2. Then F has p^m elements. Let $W_\alpha=\{1+\alpha x^2 : x \in F\}$. Let us find the number of distinct elements in W_α. For this purpose we should check that how often $1+\alpha x^2=1+\alpha y^2$. We have

$$\begin{aligned} & 1+\alpha x^2=1+\alpha y^2 \\ \Rightarrow\ & \alpha x^2=\alpha y^2 \\ \Rightarrow\ & x^2=y^2 \qquad [\because\ \alpha \neq 0] \\ \Rightarrow\ & x=\pm y. \end{aligned}$$

Thus for $x \neq 0$ we get from each pair x and $-x$ one element in W_α, and for $x=0$, we get $1 \in W_\alpha$. Consequently W_α has

$$1+\frac{p^m-1}{2}=\frac{p^m+1}{2}$$

distinct elements. Similarly we can show that $W_\beta=\{-\beta x^2 : x \in F\}$ has $\frac{p^m+1}{2}$ distinct elements. Thus W_α and W_β are subsets of F and each of them has more than half the elements of F. So the intersection of W_α and W_β cannot be empty. Let $d \in W_\alpha \cap W_\beta$.

Since $d \in W_\alpha$, therefore $d=1+\alpha a^2$ for some $a \in F$. Further since $d \in W_\beta$, therefore $d=-\beta b^2$ for some $b \in F$. Thus there exist $a, b \in F$ such that $1+\alpha a^2=-\beta b^2 \Rightarrow 1+\alpha a^2+\beta b^2=0$.

Exercises

1. (i) Define a field extension and the degree of a field extension. Find the degree and a basis of Q($\sqrt{2}, \sqrt{3}$) over Q where Q is the field of rational numbers. **[Meerut 1980, 81, 82, 84P]**

   Ans. Degree is 4, and a basis is $\{1, \sqrt{2}, \sqrt{3}, \sqrt{2}\,\sqrt{3}\}$.

   (ii) Determine the degree and a basis of the field of complex numbers over the field of real numbers. **(Meerut 1980, 82P, 83)**

2. Prove that every complex number is algebraic over the field of real numbers.
3. Find the relation between the fields Q ($\sqrt{2}$) and Q ($3+\sqrt{2}$) where Q is the field of rational numbers. **(Meerut 1977)**

   Ans. Q($\sqrt{2}$)=Q($3+\sqrt{2}$).

4. Let Q denote the field of rational numbers, K=Q($\sqrt{5}$), $L=K(\sqrt{7})$. Prove that $[L : K]=2$ and $[K :$ Q$]=2$. What do you conclude about $[L :$ Q$]$? **Ans.** $[L : Q]=4$.
5. Define splitting field of a polynomial. Show that given any non-zero polynomial $f(x)$ over a field K, then any two splitting fields of $f(x)$ over K are isomorphic. **(Meerut 1976, 77)**
6. Find the splitting field of x^4-1 over the field Q of rational numbers. **(Meerut 1978, 81, 82P)**

   Ans. Q (i).

7. Write the splitting fields of each of the following polynomials over the field Q of rational numbers.

   (i) x^2-4, (ii) x^2+4, (iii) x^2-2, (iv) x^3-1. **(Meerut 1979)**

   Ans. (i) Q, (ii) Q (i), (iii) Q($\sqrt{2}$), (iv) Q ($\sqrt{3}i$).

8. Let F be the field of rational numbers. Determine the degree of the splitting field of the polynomial x^7-1 over F.

   Ans. 6.

9. Show that the polynomials x^2+3 and x^2-x+1 have the same splitting field over F, the field of rational numbers.

10. Define splitting field of a polynomial over a field. Determine the splitting field of x^4-2 over the field of rational numbers. What is the degree of this splitting field ? **(Meerut 1978)**
Ans. Splitting field is Q $(2^{1/4}, i)$ and its degree is 8.

11. If α, β are constructible real numbers, prove that $\alpha+\beta$, $\alpha-\beta$, $\alpha\beta$ and α/β $(\beta\neq0)$ are also constructible. **(Meerut 1977)**

12. Prove that the set of all constructible real numbers form a subfield of the field of all real numbers. **(Meerut 1976)**

13. (i) All non-constructible real numbers are irrational and all irrational numbers are non-constructible. Find fault with this statement.
(ii) Give two irrational numbers, one of which is constructible and the other non-constructible (by straight edge and compass). **(Meerut 1980)**

14. Which of the following angles are constructible and which not constructible ?
10°, 15°, 1°. **(Meerut 1977, 79)**
Ans. 15° is constructible while 10° and 1° are not constructible.

15. If a field F has q elements, then F is a splitting field of x^q-x over its prime subfield. **(Meerut 1976, 79, 80)**

INDEX

A
Abelian group 49
Abel's theorem 556
Abstractly identical 128
Addition modulo *m* 77
Addition of two polynomials 327
Addition of vectors 395
Algebraic element 485
Algebraic extension 487
Algebraic structure 49
Alternating group of degree *n* 144
Annihilator 463
Associates 323
Associative operations 34
Associativity of mappings 31
Automorphisms of a group 221
Automorphism of a field 524

B
Basis of a vector space 423
Bijection 22
Binary operation 33
Boolean ring 267

C
Cayley's theorem 167
Cancellation laws in a ring 261
Cartesian product of two sets 15
Cauchy's theorem 250
Cauchy's theorem for abelian groups 249
Centre of a group 200
Characteristic of a field 293
Characteristic of a ring 291
Class equation of a group 200
Class of sets 8
Closure property 49
Co-domain of a function 19
Commutative group 49
Commutative operations 34
Commutative ring 255
Commutator subgroup of a group 239
Complementary subspaces 446
Complement of a set 11
Complexes 137
Composite of mappings 29
Composition preserving mapping 127
Composition series of a group 234
Composition table 69
Congruence of integers 83
Conjugate classes 198
Conjugate elements of a group 197
Conjugate subgroups 203
Constant function 23
Constant term 328
Constructible number 547
Content of a polynomial 382
Continuation of isomorphic mapping 504
Co-ordinates of a vector 447
Cosets 152
Cyclic groups 170
Cyclic modules 476
Cyclic permutations 98

D
Decomposition field 503
Degree of a field extension 481
Degree of a polynomial 327
Degree of an algebraic element 487
Derivative of a polynomial 516
Difference of two sets 11
Dimension of a vector space 426
Direct products of groups 244
Direct sum of spaces 443
Direct sum of submodules 473
Disjoint cycles 100
Disjoint sets 7
Disjoint subspaces 444
Distributive operations 34
Divisibility in an integral domain 321
Division algorithm for integers 82
Division algorithm for polynomials 336
Division ring 262

Domain of a function 19
Dual basis 457
Dual space 453

E

Eisenstein's criterion 391
Elementary symmetric functions 529
Elements of a set 4
Empty set 6
Endomorphism of groups 211
Equality of two functions 20
Equality of two permutations 94
Equality of two polynomials 327
Equality of two sets 5
Equivalence classes 41
Equivalence relations 40
Euclidean algorithm 339
Euclidean domains 370
Euclidean rings 370
Euler's ϕ-function 90
Euler's theorem 161
Even permutations 102
Extension of a function 23
External composition 395
External direct product of groups 244

F

Factor group 205
Factor theorem 344, 499
Family of sets 8
Fermat's theorem 162
Field 262
Field adjunction 484
Field extension 481
Field of rational functions 335
Finite dimensional vector space 424
Finite field 556
Finite integral domain 264
Finitely generated modules 476
Finitely generated vector space 424
Finite sets 4
Fixed field 525
Function 19
Fundamental theorem :
 of Galois theory 539
 on homomorphism of groups 216
 on homomorphism of modules 476
 on homomorphism of rings 359

G

Galois group 537
Gaussian integer 270
Gauss' Lemma 386
Generator of a cyclic group 170
Greatest common divisor 325
Greatest common divisor of integers 82
Greatest common divisor of two polynomials 336
Group 49
Group of automorphisms of a cyclic group 226
Group of automorphisms of a group 223
Group of permutations 96

H

Homomorphic image of a group 211
Homomorphism of groups 211
Homomorphism of modules 474
Homomorphism of rings 356
Homomorphism of vector spaces 433

Ideals 306
Identity element 35
Identity mapping 23
Identity permutation 95
Image of a subset 20
Imbedding of a ring into another ring 296
Improper divisors 324, 335
Improper subgroups 138
Improper subset 7
Index of a subgroup 157
Index set 12
Induced composition 137
Infinite sets 4
infinite order 113
Inner automorphisms 224
Integral domain 261
Integral multiples of an element of a group 112

Integral multiples of an element of a ring 256
Integral powers of an element of a group 111
Internal composition 395
Internal direct products of groups 246
Intersection of sets 10
Intersection of subgroups 145
Intersection of subrings 287
Into mappings 21
Invariant subgroups 204
Inverse function 24
Inverse image 23
Inverse relations 37
Inversible elements 35
Irreducible elements 325
Irreducible module 472
Irreducible polynomials 335
Isomorphism of groups 126
Isomorphism of rings 282
Isomorphism of vector spaces 435

J

Jorden Holder theorem 235

K

Kernel of a group homomorphism 213
Kernel of a module homomorphism 474
Kernel of a ring homomorphism 357
Kernel of a vector space homomorphism 434
Kronecker delta 455

L

Lagrange's theorem 159
Leading coefficient 328
Leading term 328
Left cancellation law 57
Left coset 152
Left coset decomposition of a group 159
Left ideal 306
Left identity 58
Left inverse 58
Left module 468
Linear combination 410
[illegible] of vectors 413
Linear functional 449
Linear independence of vectors 414
Linear operator 449
Linear span 410
Linear sum of two subspaces 412
Linear transformation 433

M

Many-one mappings 21
Mapping 19
Maximal ideal 361
Maximal subgroups 233
Members of a set 4
Minimal polynomial 486
Module 468
Monic polynomial 335
Multiple root 499
Multiplication modulo *p* 77
Multiplication of two polynomials 327

N

Non-empty set 6
Normal extension of a field 529
Normalizer of an element of a group 198
Normalizer of a subgroup of a group 203
Normal series of a group 236
Normal subgroup 188
Null ideal 307
Null set 6

O

Odd permutations 102
One-one mappings 21
One-to-one correspondence 22
Onto mappings 21
Operators 20
Ordered basis 447
Ordered field 294
Ordered integral domain 294
Ordered pair 15
Order of a group 50
Order of an element of a group 113
Order relations in an integral domain 295
Outer automorphism 224

P
Partial order relation 46
Partitions 43
Perfect field 523
Permutations 93
Polynomial equations 344
Polynomials 326
Polynomials over a field 333
Power set 8
Prime elements of an integral domain 325
Prime fields 345
Prime ideal 365
Prime integers 83
Primitive polynomials 383
Principal ideal 319
Principal ideal ring 320
Principle of trichotomy 294
Product of mappings 29
Product of permutations 95
Proper divisors 324, 335
Proper ideals 307
Proper subgroups 138
Proper subset 7

Q
Quaternion group 74
Quotient field 299
Quotient groups 205
Quotient modules 476
Quotient rings 354
Quotient set 45
Quotient space 442

R
Range of a function 19
Reflexive relations 38
Reflexivity of vector spaces 460
Relation 36
Relatively prime elements 325
Relatively prime integers 82
Relatively prime polynomials 336
Remainder theorem 344, 498
Residue classes of integers modulo *m* 84
Restriction of a function 23
Right cancellation law 57
Right coset 152
Right coset decomposition of a group 155
Right ideal 306
Right module 468
Ring 254
Ring of endomorphisms of an abelian group 348
Ring of Gaussian integers 271
Ring of integers 257
Ring of polynomials 330
Ring of residue classes 354
Ring without zero divisors 260
Ring with unity 255
Root of a polynomial 485, 498

S
Scalar multiplication 395
Scalars 395
Second dual space 459
Second law of isomorphism 230
Semi-group 49
Separable element 523
Separable extension 523
Set 3
Simple field extension 485
Simple group 188
Simple ring 307
Simple root 499
Singleton 6
Skew field 262
Solvable by radicals 552
Solvable group 236
Splitting field 503
Standard basis 4?3
Subfields 290
Subgroup 137
Subgroup generated by a subset of a group 184
Submodules 471
Subrings 286
Subsets 6
Superset 7
Sylow's theorem 251
Sylow subgroup 251
Symmetric difference 12
Symmetric group of degree *n* 98
Symmetric relations 39

T

Tautology 2
Third law of isomorphism 231
Transcendental element 486
Transcendental extension 487
Transformations 20
Transitive relations 40
Transposition 99

U

Union of sets 9
Unique factorization domain 340
Unique factorization theorem for polynomials 342
Unique factorization theorem for Euclidean rings 378
Uniqueness of the splitting field 512
Unit ideal 307
Units 322
Universal set 9

V

Vectors 395
Vector space 395
Vector subspace 403
Venn diagrams 9
Void set 6

Z

Zero divisors 260
Zero element of a ring 254
Zero order element of a group 113
Zero polynomial 327
Zero ring 257
Zero of a polynomial 344
Zero subspace 406
Zero vector 396